5-31-2017 12:00 AM

Characterizing the Interaction Between Human Adenovirus E1A and the Transcriptional Repressor BS69

Ali Zhang
The University of Western Ontario

Supervisor
Dr. Joe Mymryk
The University of Western Ontario

Graduate Program in Microbiology and Immunology
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

Abstract

Protein products of the Early Region 1A (E1A) gene in human adenovirus 5 (HAdV-5) are the first viral proteins expressed upon adenovirus infection. E1A disrupts many cellular physiological events by binding to and regulating an impressive number of host factors. Of particular interest is BS69, a repressor of E1A transactivation. Due to the strong interaction observed between E1A and BS69, I hypothesize that these two proteins function together to disrupt gene expression within an infected cell.

Using *in silico* modelling and a series of yeast two-hybrid assays, I determined that residues 112-119 of HAdV-5 E1A is the minimal interacting region for BS69. This interaction is conserved in HAdV-5, 9, and 12 from species C, D, and A. Furthermore, I found that the MYND domain of BS69 is both necessary and sufficient to interact with and inhibit E1A-mediated transactivation in a mechanism dependent on the fidelity of the PXLXP motif and the adenovirus species. Therefore, I have found that BS69 physically interacts with E1A to disrupt gene transcription in mammalian cells. Future studies will reveal the effects of the E1A-BS69 interaction on viral growth and regulation of viral and host genes.

Keywords

Human Adenovirus, E1A, BS69, ZMYND11, Yeast Two-Hybrid, Co-Immunoprecipitation, Peptide Array, Transcription, Luciferase Assay, Protein-Protein Interaction

Acknowledgments

I would first like to thank my supervisor, Dr. Joe Mymryk, for the mentorship that he has provided me with for past three years. This has been an amazing experience and I am truly grateful to have had the privilege of being a member of his lab. The learning environment here is outstanding, and is a direct product of Dr. Mymryk's wisdom, kindness, and humility. At the Mymryk lab, every day is an exciting opportunity to learn something new. I will dearly miss the lessons on automotive maintenance, computer hardware, and home ownership.

I would also like to thank my committee members Dr. Jim Koropatnick and Dr. Joe Torchia, and Dr. Jimmy Dikeakos, Dr. Kelly Summers, and Dr. Ewa Cairns from the Department of Microbiology and Immunology for their insight and guidance throughout graduate school. Your expertise was greatly appreciated, and was instrumental to the success of this project. Thank you all for the advice and kind words of encouragement.

A large portion of this project would not have been possible without the help of Dr. John Barrett, Gloria Thomson, and Mike Qin. Dr. John Barrett, for his technical expertise and advice; Gloria Thomson, for her enthusiastic willingness to lend a hand; and Mike Qin, without whom the peptide array experiments would not have been possible.

Graduate school can be stressful, so I am extremely fortunate to have been surrounded by a wonderful group of friends and who have made it better. I highly enjoyed watching Star Wars, having exotic lab lunches, and destroying online noobs with you. Thank you to the past and present lab members: Dr. Mike Cohen, Cason King, Tanner Tessier, Steve Gameiro, Kristianne Galpin, Martin Prusinkiewicz, Jessica Hill, Nicole Pinto, Kara Ruicci, Farhad Ghasemi, and Morgan Black. You have all made a positive impact on my life, and I am very grateful for that.

Thank you, Jean, for everything you have given me throughout the years. You have helped me learn and grow into the person I am now. Thank you for pushing me in the right direction and for always being by my side.

Finally, I would like to thank my mom, dad, and younger brother Adam for their endless love and encouragement. I could not have accomplished this without all of you, and for that, I owe you all my deepest gratitude.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Ad Pol	Adenovirus DNA Dependent DNA Polymerase
Amp	Ampicillin
AmpR	Ampicillin Resistance Gene
AP2	Activating Protein 2
ATF	Activating Transcription Factor
ATM	Ataxia-Telangiectasia Mutated
ATR	ATM- and Rad3-related protein
Bcl-2	B-cell Lymphoma 2 protein
BME	β-mercaptoethanol
BMP	Bone Morphogenetic Protein
BRAM1	Bone Morphogenetic Protein Receptor Associated Molecule 1
BSA	Bovine Serum Albumin
CAF1	Chromatin Assembly Factor 1
Cam	Chloramphenicol
CAR	Coxsackie Adenovirus Receptor
CCR4	C-C Motif Chemokine Receptor 4
ChIP-seq	Chromatin Immunoprecipitation Sequencing
Co-IP	Co-immunoprecipitation
CR	Conserved Region
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DBD	DNA Binding Domain
DCM	Dichloromethane
ddH2O	Double Distilled Water
DDR	DNA Damage Response
DEAF1	Deformed Epidermal Autoregulatory Factor 1 homolog
DMEM	Dulbecco's Modified Eagle Medium
DMF	Dimethylformamide
DNABP	DNA Binding Protein
EBNA2	Epstein Barr Nuclear Antigen 2
EBV	Epstein Barr Virus
ECL	Enhanced Chemiluminescence

EFTUD2	Elongation Factor Tu GTP-Binding Domain-Containing Protein 2
ER	Endoplasmic Reticulum
FBS	Fetal Bovine Serum
FIP	14.7K-Interacting Protein
G6PD	Glucose-6-Phosphate Dehydrogenase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green Fluorescent Protein
GST	Glutathione S-transferase
HEK293	Human Embryonic Kidney 293
HF	High Fidelity
HRP	Horseradish Peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
KanR	Kanamycin Resistance
L	Major Late Unit
LB	Lysogeny Broth
LiAc	Lithium Acetate
MAPK	Mitogen-Activated Protein Kinase
MCS	Multiple Cloning Site
MES	2-[N-morpholino]ethanesulfonic acid
mTOR	Mechanistic Target of Rapamycin
MYND	Myeloid, Nervy, and DEAF-1
N-CoR2	Nuclear Receptor Co-Repressor 2
NMR	Nuclear Magnetic Resonance
ONPG	Ortho-Nitrophenyl-β-Galactoside
ORF	Open Reading Frame
PA	Peptide Array
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol 3350
PHD	Plant Homeodomain
PI3K	Phosphatidylinositol 3-Kinase

PIKK	Phosphatidylinositol 3-Kinase-related Kinase
PKA	Protein Kinase A
PKR	Protein Kinase R
PML	Promyelocytic Leukemia
PMSF	Phenylmethylsulfonyl Fluoride
PNK	Polynucleotide Kinase
PNN	Pinin
PONDR	Predictor of Naturally Disordered Regions
PP2A	Protein Phosphatase 2
pRb	Retinoblastoma Protein
PRPF8	Pre-mRNA Processing Splicing Factor 8
pTP	Terminal Protein
PVDF	Polyvinylidene Fluoride
qPCR	Quantitative Polymerase Chain Reaction
Ras	Rat Sarcoma
RCF	Relative Centrifugal Force
RID	Receptor Internalization and Degradation
RNAPII	RNA Polymerase II
SDS	Sodium Dodecyl Sulfate
SH3	Src Homology 3
shRNA	Short Hairpin RNA
siRNA	Silencing RNA
skNAC	Skeletal α-Nascent Polypeptide-Associated Complex
SLiM	Short Linear Interaction Motif
SNRNP200	U5 Small Nuclear Ribonucleoprotein 200kDa Helicase
SRSF1	Serine/Arginine-rich Splicing Factor 1
SS-DNA	Salmon Sperm DNA
STING	Stimulator of Interferon Genes
SUMO	Small Ubiquitin-like Moiety
TBS-T	Tris Buffered Saline with Tween-20
TFA	Trifluoroacetic Acid
TIPS	Triisopropylsilane

TNF	Tumour Necrosis Factor
TNFR1	Tumor Necrosis Factor Receptor 1
TR	Thyroid Hormone Receptor
TRAILR	Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand Receptor
U	U Exon
UAS	Upstream Activation Sequence
UBC9	Ubiquitin Conjugase 9
USF	Upstream Stimulatory Factor
UXP	U Exon Protein
VA RNA	Virus-Associated RNA
WB	Western Blot
WT	Wildtype
YEP	Yeast Extract Peptone
ZHX	Zinc Fingers and Homeoboxes

1 Introduction

Viruses are obligate intracellular parasites. It is not known how viruses originated, and it is still unclear if viruses are considered living organisms. At the most basic level, viruses are composed of a DNA or RNA genome encased within a proteinaceous capsid. The viral genome encodes structural and functional proteins to help with infection and replication. Some “deluxe” viruses are also covered by a lipid envelope. Viruses exist in all shapes and sizes, and infect organisms from all domains of life. They range from the tiny Circovirus, which can have a genome of less than 1800 nucleotides and a capsid of 15nm in diameter, to the humongous Pithovirus, which can have a genome of over 600 000 nucleotides and a capsid of up to 1500nm in length. Viruses are diverse and fascinating, and are intricately intertwined with the evolution of their host.

Viruses were discovered in the early 1900s and have since played a pivotal role in the history of healthcare and biomedical research. Viruses cause many devastating diseases including smallpox and polio, and are thus an intensively studied topic. Because viruses have co-evolved extensively with their hosts, virus research has also taught us many lessons in cell biology. One specific example is the identification of mRNA splicing, which is a crucial cellular process discovered through examination of the adenovirus genome. With a relatively small genome and limited coding capacity, viruses also have the challenging task of overcoming host immunity and repurposing the cell to become conducive for viral replication. Thus, targets of viral proteins are likely critical cogs in the molecular machine.

This project is focused on better understanding the biology of human adenoviruses and how they interact with the host. Specifically, we examine the components and effects resulting from the interaction between the adenovirus E1A protein and the host transcription regulator, BS69. From this project, we hope to learn more about viral-host interactions and to better characterize this mechanism of viral gene regulation.

1.1 Adenoviruses

1.1.1 Discovery and Classification of Adenoviruses

Adenoviruses were discovered by two separate groups in search of a causative agent of respiratory disease in 1953 and 1954 (Hilleman & Werner, 1954; Rowe et al., 1953). Rowe *et al*. found a cytopathic agent isolated from human adenoid tissue that caused tissue degeneration in culture, while Hilleman and Werner isolated an agent with similar cytopathic effects from throat washings of an army recruit with primary atypical pneumonia (Hilleman & Werner, 1954; Rowe et al., 1953). These cytopathic agents were later determined to be viruses, and were named adenovirus in 1956 after the source, human adenoid tissue, from which the original isolate was found (Enders et al., 1956; Huebner et al., 1954).

Adenoviruses are classified under the family *Adenoviridae* and branch into five genera: *Atadenovirus, Aviadenovirus, Ichtadenovirus, Mastadenovirus,* and *Saidenovirus*. Human adenoviruses (HAdV), along with adenoviruses that infect other mammals and vertebrates, are classified under the *Mastadenovirus* genus. Mastadenoviruses are further categorized into 7 species, adenovirus species A through G, based on biological properties such as hemagglutination groups, oncogenic potential, and genome sequence homology (Table 1.1). Furthermore, there are currently 57 accepted human adenovirus types (HAdV-1 to 57) distributed amongst these 7 adenovirus species. The newest species is species G, with HAdV-52 being the only member, while the newest type is HAdV-57 (Jones et al., 2007; Walsh et al., 2011). Human adenovirus types were historically distinguished based on resistance to neutralization by antisera against other adenovirus types. However, more recent techniques focus on protein or nucleotide sequence alignment in the main type-specific epitopes, namely loop 1 and loop 2 of the hexon protein, involved in virus neutralization tests (Madisch et al., 2005). A ≥1.2% genetic divergence of loop 2 from the closest adenovirus prototype, or ≥2.4% genetic divergence in combination with ≥4.2% amino acid divergence of loop 1 from the closest prototype is necessary to support the identification of a new adenovirus type (Madisch et al., 2005). Other methods of adenovirus typing include sequence comparison of fiber knobs and hemagglutinin inhibition testing (Madisch et al., 2005).

Table 1.1. Properties of the seven human adenovirus species. Table showing hemagglutination group, example HAdV types, tumorigenicity in animals, transformation in tissue culture, genome percent GC, and infection/disease type of the seven HAdV species. HAdV-5 of species C is the most studied type, and the focus of this thesis. Table was adapted from (Berk, 2007), infection/disease type from (Ghebremedhin, 2014), and information of HAdV species G from (Jones et al., 2007).

Species	Hemagglutination Group	Adenovirus Types	Tumours in Animals	Transformation in Tissue Culture	Percentage of GC in DNA	Type of Infection / Disease
A	IV (little or no agglutination	12, 18, 31	High	+	48-49	Gastrointestinal, respiratory, urinary
B	I (complete agglutination of monkey erythrocytes)	3, 7, 11, 14, 16, 21, 34, 35, 50	Moderate	+	50-52	Keratoconjunctivitis, gastrointestinal, respiratory, urinary
C	III (partial agglutination of rat erythrocytes)	1, 2, 5, 6	Low or none	+	57-59	Respiratory, gastrointestinal, urinary
D	II (complete agglutination of rat erythrocytes)	8, 9, 10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, 42-49, 51	Low or none	+	57-61	Keratoconjunctivitis, gastrointestinal
E	III	4	Low or none	+	57-59	Keratoconjunctivitis, respiratory
F	III	40, 41	Unknown	Unknown	Unknown	Gastrointestinal
G	Unknown	52	Unknown	Unknown	55	Gastrointestinal

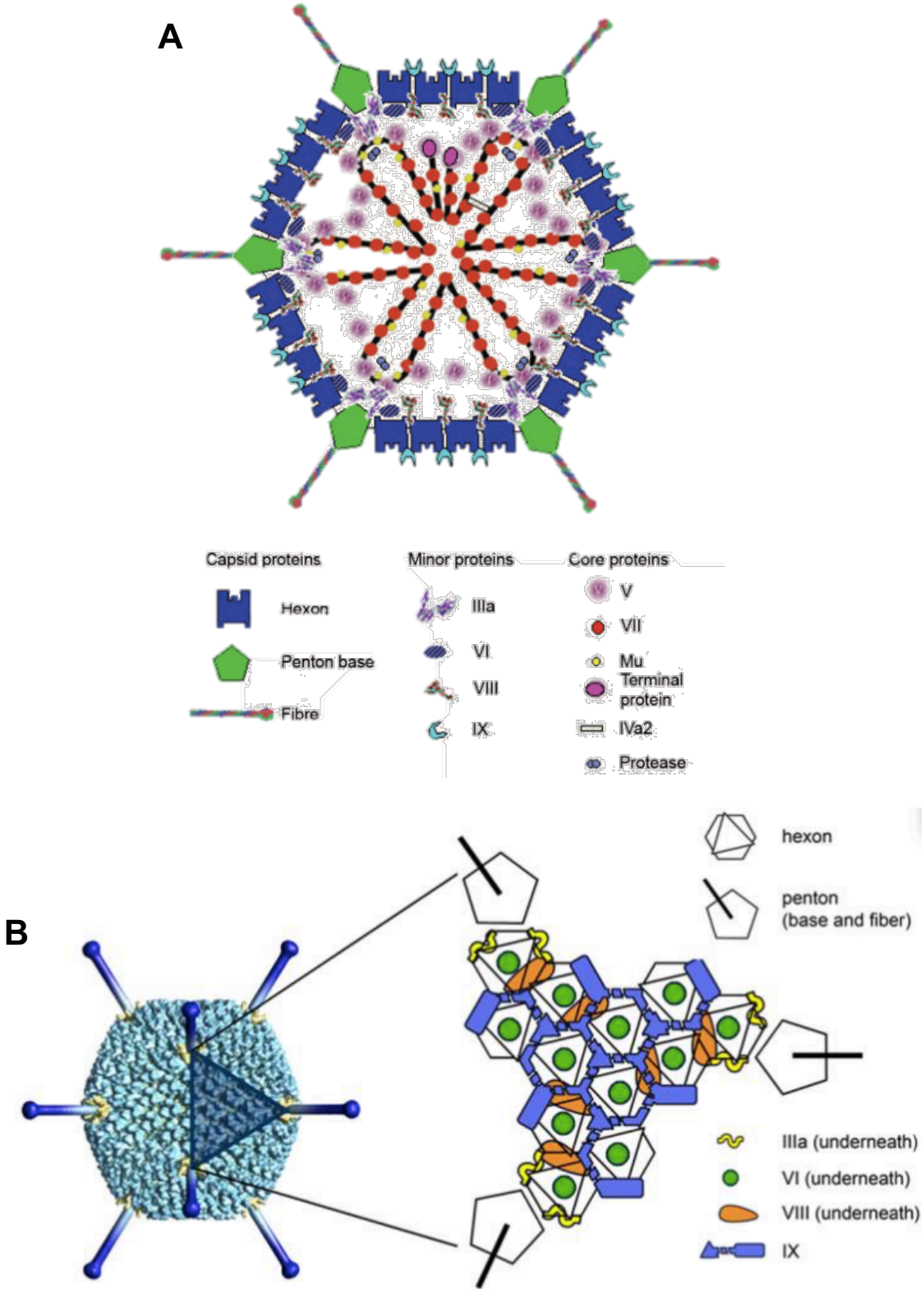
1.1.2 Human Adenovirus Virion Structure

HAdV are non-enveloped viruses that are 80-110nm in diameter and encased by an icosahedral nucleocapsid composed of 252 subunits (Figure 1.1). 240 of these subunits are hexon capsomeres, while 12 are penton capsomeres. Hexon capsomeres are homotrimers that include up to 9 hypervariable loops that make up one of the major adenovirus type-specific antigenic regions (Russell, 2009). These capsomeres are surrounded by 6 other penton or hexon subunits and form the 20 faces of the icosahedral nucleocapsid (San Martín, 2012). Penton capsomeres are homopentamers and include an arginine-glycine-aspartic acid, or RGD, motif used to facilitate virus internalization via attachment to cellular integrins (Wickham et al., 1993). Penton capsomeres are surrounded by 5 other hexon subunits and form the 12 vertices of the icosahedral nucleocapsid. A homotrimeric fiber protein is non-covalently bound to each penton capsomere through a conserved N-terminal FNPVYPY sequence on the fiber protein (Tarassishin et al., 2000). The flexibility and the length of the fiber shaft varies greatly, as the fiber is composed of 3 to 23 pseudorepeats with 15 to 20 residues per repeat depending on the adenovirus type (van Raaij et al., 1999). The fiber shaft is also glycosylated, but the function of these modifications is unknown (Cauet et al., 2005). The fiber shaft terminates with a globular knob domain used to interact with receptors on host cells to facilitate initial virus attachment. The fiber protein targets for most HAdV species have yet to be identified, but the fiber protein from HAdV species C interacts with the Coxsackie adenovirus receptor (CAR) (Bewley et al., 1999).

1.1.3 Human Adenovirus Disease and Tropism

HAdV are ubiquitous in humans and can infect multiple organ systems; however, these infections are usually mild, self-limiting, and may cause illnesses including pharyngitis, gastroenteritis, or conjunctivitis. HAdV infections are more prevalent and may cause more severe diseases, such as pneumonia, fulminant hepatitis, and/or encephalitis in children and immunocompromised patients (Krilov, 2005; Walls et al., 2003). The type of disease caused by this virus is also dependent on the HAdV species; for example, species D

Figure 1.1. Cartoon depiction of the adenovirus virus particle. A) Cross-section diagram showing the arrangement of the various adenovirus structural proteins. Green pentagons depict the penton capsomeres, while the blue polygons depict the hexon capsomeres. The genome is depicted with the black line depicted in the shape of an asterisk inside of the capsid. Adapted from (Russell, 2009). B) Diagram showing one of twenty faces of the icosahedral virus capsid. Pentons are shown as pentagons, hexons are shown in hexagons, and various other structural proteins labelled in the legend. Adapted from (San Martín, 2012).


A
Capsid proteins
Hexon
Penton base
Fibre
Minor proteins
IIIa
VI
VIII
IX
Core proteins
V
VII
Mu
Terminal protein
IVa2
Protease
B
hexon
penton
(base and fiber)
IIIa (underneath)
VI (underneath)
VIII (underneath)
IX

typically cause conjunctivitis, species E causes respiratory infections, and species F and G typically cause gastroenteritis (Table 1.1) (Chang et al., 2008; Sambursky et al., 2007; Walls et al., 2003). Some HAdV types have also been associated with severe disease outbreaks; for example, HAdV-14 caused a fatal pneumonia outbreak in patients living in residential care facilities and recruits in military training centers in the United States in 2005 (Lewis et al., 2009). HAdV-36 has also been linked to obesity in humans, and this virus has been experimentally verified to caused increased adiposity in animal models as well (Atkinson et al., 2005; Atkinson, 2007; Dhurandhar et al., 2000).

1.1.4 Human Adenovirus Oncogenesis

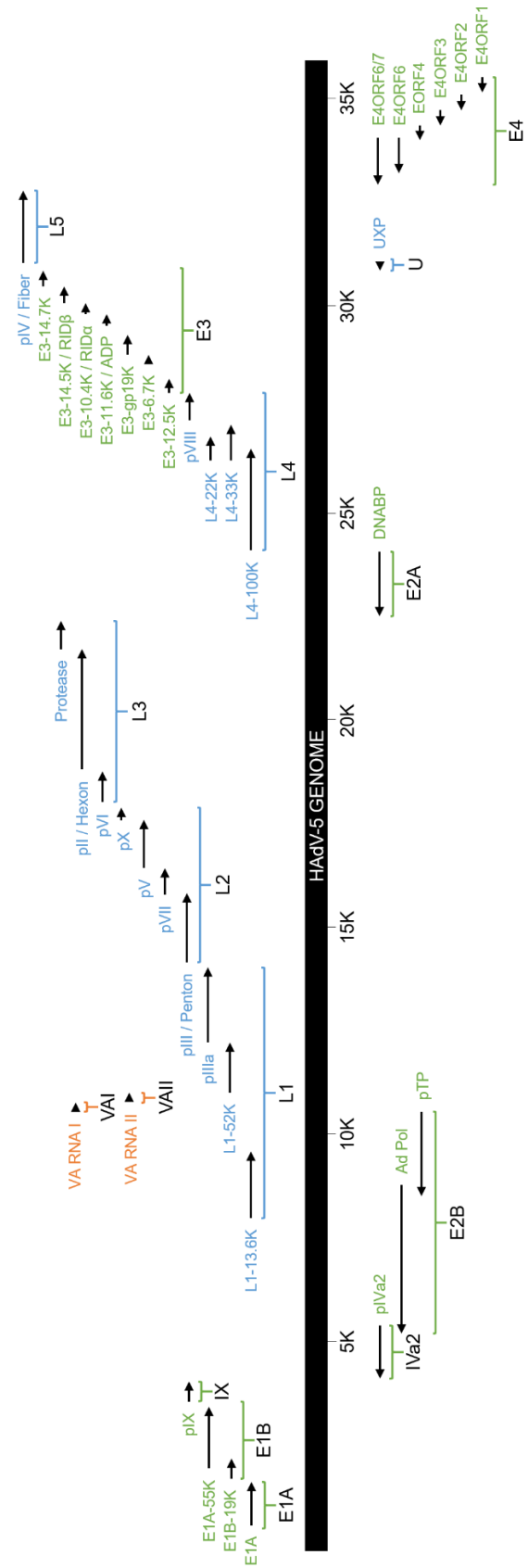
Although adenoviruses are classified as small DNA tumour viruses, HAdV do not cause cancer in humans (Green et al., 1980; Mackey et al., 1976). These viruses are, however, tumourigenic in some rodent species and can transform primary rodent cells in culture (Gallimore, 1972; Trentin et al., 1962). The mechanism of transformation by adenoviruses require both the viral E1A and E1B proteins (Sherr & McCormick, 2002; Van den Elsen et al., 1983). E1A abrogates the activity of retinoblastoma protein (pRb) and its family members p107 and p130, which are repressors of the E2F family of transcription factors (Cobrinik, 2005; Frolov & Dyson, 2004). The E2F family of transcription factors is largely responsible for controlling cell cycle progression from the G1 to S phase (Dyson, 1998). By binding to and displacing pRb and its family members from E2F, E1A effectively helps circumvent this cell cycle checkpoint to mediate aberrant cell division (Bagchi et al., 1990; Zamanian & La Thangue, 1992). Additionally, E1A blocks the activity of cyclin dependent kinase inhibitors p16INK4, p21, and p27kip1 to further promote cell cycle progression from the G1 to S phase (Alevizopoulos et al., 1998, 2000; Chattopadhyay et al., 2001; Mal et al., 1996). Successful cell transformation cannot be accomplished by E1A alone, as E1A-mediated effects on the cell also cause stabilization of p53 and subsequent p53-mediated cell apoptosis (Lowe & Ruley, 1993). Thus, the E1B-55K/E4ORF6 effects of p53 inhibition and degradation, discussed in the next section, must also be present for successful cell transformation (Cathomen & Weitzman, 2000; Van den Elsen et al., 1983).

1.1.5 Human Adenovirus Genome

The adenovirus nucleocapsid contains a 36-38kb linear double-stranded DNA genome, which encodes 34-38 proteins from both the forward and reverse strands (Figure 1.2). A 55kDa terminal protein (pTP) is covalently linked to each 5' end of adenovirus genome, and function as primers for viral replication (Challberg et al., 1980). The adenovirus genome can be categorized into transcription units that are expressed either early, delayed, or late into the infection cycle. The late phase of an adenovirus infection is defined by the commencement of adenoviral genome replication, which occurs approximately 6 hours post infection under ideal conditions in tissue culture systems. Early transcription units include E1A, E1B, E2B, E3, and E4; delayed units include E2A, IX, and IVa2; and late units consist of the major late unit (L) and U exon (U) (Fessler & Young, 1998; Tollefson et al., 2007). These transcriptional units occupy both strands of the adenoviral genome, with E1A, E1B, E3, IX, and L on the rightward transcription strand, while E2A, E2B, E4, IVa2, and U are on the leftward transcription strand (Figure 1.2). Cellular RNA polymerase II (RNAPII) transcribes all the aforementioned regions (Berk, 1986). Additionally, depending on the virus type, adenoviruses also encode one or two non-coding virus-associated RNA (VA RNA); these VA RNAs are approximately 160 nucleotides in length, localized to the cytoplasm, transcribed by RNA polymerase III, and function to antagonize host antiviral responses by blocking the interferon-induced activation of protein kinase R (PKR) (Mathews & Shenk, 1991).

Adenovirus early genes encode for proteins with a variety of functions including regulation of viral gene expression, replication of the viral genome, and repression of host antiviral responses. E1A is the first region expressed during an adenovirus infection, and is located on the leftmost portion of the genome (Nevins et al., 1979). The E1A protein is the main topic of this thesis, and will be discussed in more detail in later sections. Briefly, key functions of E1A include driving quiescent cells into cell cycle by inducing the disassembly of the E2F1-pRb complex, and transactivation of other adenovirus early genes (Berk, 1986; DeCaprio, 2009). In HAdV-5, the E1B region codes for two proteins: E1B-19K and E1B-55K. E1B-19K and E1B-55K block host cell apoptosis by mimicking the anti-apoptotic B-cell lymphoma 2 protein (Bcl-2), and by forming a complex with E6ORF6 to facilitate the

Figure 1.2. Map of transcripts in the HAdV-5 genome. The HAdV-5 genome is approximately 36kb long, as shown by the ruler under the black bar. Arrows depict the direction of individual transcripts, with the protein product labelled beside the arrow. The transcriptional units are labelled below the individual transcripts. Early and delayed genes are in green, while late genes are labelled in blue. Orange transcripts depict the HAdV-5 VA RNAs. Reference genome was obtained from NCBI, with the RefSeq code NC_001405.1.


HAdV-5 GENOME
5K
10K
15K
20K
25K
30K
35K
E1A
E1A
E1B-19K
E1A-55K
E1B
pIX
IX
E2B
pTP
Ad Pol
IVa2
pIVa2
L1
L1-13.6K
L1-52K
pIIIa
pIII / Penton
L2
pVII
pV
pX
pVI
pII / Hexon
Protease
L3
VA RNA II
VAII
VA RNA I
VAI
L4
L4-100K
L4-33K
L4-22K
pVIII
E3-12.5K
E3-6.7K
E3-gp19K
E3-11.6K / ADP
E3-10.4K / RIDα
E3-14.5K / RIDβ
E3-14.7K
pIV / Fiber
L5
E2A
DNABP
U
UXP
E4ORF1
E4ORF2
E4ORF3
E4ORF4
E4ORF6
E4ORF6/7
E4

degradation of the p53 tumour suppressor protein respectively (White, 2001). E1B-55K also plays a crucial role in the export of viral mRNA from the nucleus of host cells to help with viral protein production (Gonzalez & Flint, 2002). The E2B region encodes two proteins necessary for adenovirus genome replication: pTP, which is covalently linked to each 5' end of the adenovirus genome to function as primers for viral DNA synthesis, and the adenovirus DNA dependent DNA polymerase (Ad Pol), which is used to replicate the viral genome (Challberg et al., 1980; Field et al., 1984).

In species C adenoviruses, the E3 region encodes seven immunomodulatory and cell lysis proteins. These proteins include E3-12.5K, E3-6.7K, E3-gp19K, adenovirus death protein (ADP, also known as E3-11.6K), receptor internalization and degradation protein α (RIDα, also known as E3-10.4K), receptor internalization and degradation (RID) protein β (RIDβ, also known as E3-14.5K), and E3-14.7K (Lichtenstein et al., 2004). E3-12.5K was discovered in 1992 and is highly conserved between HAdV-2, 3, and 5; however, the function of this protein is currently unknown (Hawkins & Wold, 1992). E3-6.7K is a type III transmembrane protein that inhibits host cell apoptosis by interacting with the RID complex (discussed below), maintaining Ca^{2+} concentrations in the endoplasmic reticulum (ER), and reducing tumour necrosis factor (TNF)-mediated release of arachidonic acid (Elsing & Burgert, 1998; Moise et al., 2002). E3-gp19K is a type I transmembrane glycoprotein that facilitates adenovirus immune evasion by inhibiting the function of major histocompatibility complex class I (MHC I) (Burgert et al., 1987; Wold et al., 1985). Immune evasion by E3-gp19K is achieved through two known mechanisms that result in delayed MHC I maturation: retaining MHC I in the ER and inhibiting the function of tapasin, which is a protein that facilitates MHC I peptide loading (Bennett et al., 1999; Burgert et al., 1987). E3-11.6K, or ADP, is unique in that although the gene is located within an early transcription unit, the protein is produced predominantly during the late stages of infection (Tollefson et al., 1992). Like E3-6.7K, ADP is also a type III membrane protein (Tollefson et al., 1996). ADP facilitates lysis of adenovirus-infected cells and subsequent release of virus particles (Tollefson et al., 1996). RIDα and RIDβ form the RID complex, which functions to block apoptosis of the host cell by downregulating proapoptotic cell surface receptors such as Fas, TNF-related apoptosis-inducing ligand receptor 1 and 2 (TRAILR1 and TRAILR2) and TNF receptor 1 (TNFR1) through receptor

internalization and subsequent lysosomal degradation (Benedict et al., 2001; Elsing & Burgert, 1998; Friedman & Horwitz, 2002). The downregulation of TRAILR1 and TRAILR2 also require an interaction between E3-6.7K and the RID complex. Finally, E3-14.7K is nonamer that antagonizes TNF-induced apoptosis (Gooding et al., 1988, 1990; Kim & Foster, 2002). The mechanism behind the antagonism of TNF-induced apoptosis is unknown, but may be related to the interaction between E3-14.7K and the 14.7K-interacting protein (FIP) family of proteins (Li et al., 1999, 1997).

The E4 transcriptional unit is located on the right end of the adenovirus genome (Figure 1.2). In species C adenoviruses, the E4 transcriptional unit encodes several proteins named after the order of open reading frames (ORF) from which they originate: E4ORF1 through E4ORF4, E4ORF6, and E4ORF6/7 (Täuber & Dobner, 2001). E4ORF1 assists in increasing viral replication by dysregulating host cell metabolism and replication through several mechanisms. Firstly, E4ORF1, in a E4ORF4-dependent mechanism, activates the mechanistic target of rapamycin (mTOR) pathway by activating phosphatidylinositol 3-kinase (PI3K) (Frese et al., 2003; O'Shea et al., 2005). E4ORF1 also changes host cell metabolism in a MYC-dependent mechanism by promoting glycolysis and shunting the resulting metabolites to nucleotide biosynthesis pathways (Thai et al., 2014). Interestingly, E4ORF1 also plays a role in obesity caused by HAdV-36 infections (Sohrab et al., 2017). E4ORF2 is localized to the cytoplasm, but the function of the protein is not known (Dix & Leppard, 1995; Thomas et al., 2001). E4ORF3 is involved in the reorganization of promyelocytic leukemia (PML) oncogenic domains and inactivation of p53 by inducing heterochromatin formation at p53 target promoters (Evans & Hearing, 2003; Soria et al., 2010). The mechanism of p53 inactivation by E4ORF3 is separate and independent from that of E1B (Soria et al., 2010). E4ORF4 inhibits the host DNA damage response (DDR) in several ways. With its cellular binding partner protein phosphatase 2 (PP2A), E4ORF4 reduces phosphorylation of substrates of the phosphatidylinositol 3-kinase-related kinase (PIKK) family including ataxia-telangiectasia mutated (ATM) and ATM- and Rad3-related protein (ATR), resulting in inhibition of various DDR pathways (Brestovitsky et al., 2016). E4ORF4 is also able to mislocalize various sensor complexes required to activate ATM and ATR, thus preventing downstream DDR signaling (Carson et al., 2009). E4ORF6 is yet another adenovirus protein that blocks p53 activity via one of two mechanisms:

E4ORF6 by itself can inhibit the interaction between p53 and its transcriptional coactivator, TAFII31; or as discussed previously, E4ORF6 forms a complex with E1B-55K to ubiquitinate p53, leading to p53 degradation (Dobner et al., 1996; Lu & Levine, 1995; Querido et al., 2001). Having so many distinct and redundant avenues of attacking the "guardian of the genome" truly illustrates the importance of dismantling the p53 pathway in a successful and productive adenovirus infection. The E4ORF6/E1B-55K complex is involved in many other processes as well, including viral DNA replication, nuclear export of late viral mRNA, and inactivation of DDR (Bridge & Ketner, 1990; Gonzalez & Flint, 2002; Stracker et al., 2002). Lastly, E4ORF6/7 is a protein that is encoded by DNA from both ORFs 6 and 7. This protein forms a complex with the E2F transcription factor, recruits it to the E2 early promotor, and induces transactivation of E2 genes (Marton et al., 1990; Obert et al., 1994).

The adenovirus delayed transcription units include E2A, IX, and IVa2. E2A encodes a DNA binding protein (DNABP) involved in viral DNA replication (Lindenbaum et al., 1986). This protein binds to single stranded DNA and stimulates Ad Pol to commence DNA synthesis (Lindenbaum et al., 1986). The IX transcription unit encodes pIX, a minor "cement" protein present on the exterior surface of the adenovirus capsid, used to link other structural proteins together and with the viral core (Liu et al., 2010; Reddy et al., 2010). Lastly, IVa2 encodes pIVa2, which is a viral packaging protein that functions as an ATPase; pIVa2 forms a complex with other late proteins and the E2A DNABP to facilitate loading of the adenovirus genome into an assembled capsid (Ahi et al., 2013; Christensen et al., 2008; Ostapchuk & Hearing, 2008).

Adenovirus late proteins are encoded by the major late transcription unit and the U exon. The major late transcription unit is further divided into L1-L5 transcription units. The U exon protein (UXP) encoded by the U exon is 24kDa in size, localized to the nucleus and viral replication centers, and may be involved in adenovirus DNA replication (Tollefson et al., 2007). Proteins encoded by the major late transcriptional unit function mostly as adenovirus structural proteins or proteins involved in the viral packaging process (Table 1.2). Notably, the penton subunit, hexon subunit, and fiber are encoded on L2, L3, and L5

Table 1.2. Proteins encoded by the major late transcription unit. Table showing details about the proteins encoded in the L1-L5 transcription units, presence in empty or mature capsids, location, and function(s). Adapted from (Ahi & Mittal, 2016), while information about L4-100K is adapted from (Cepko & Sharp, 1982).

Transcription Unit	Adenovirus Protein	Presence in Empty or Mature Capsids	Location	Function(s)
L1	L1-13.6K	Unknown	Unknown	Unknown
	L1-52K	Empty	Interior Capsid	Packaging
	pIIIa	Both	Exterior Capsid, Interior Vertex	Cement
L2	pIII (Penton)	Both	Exterior Vertex	Structure, Viral Entry
	pVII	Mature	Core	Core Condensation
	pV	Mature	Core	Core Condensation
	pX	Mature	Core	Core Condensation
L3	pVI	Both	Interior Vertex	Cement, Viral Entry, Protein Trafficking, Early Gene Expression
	pII (Hexon)	Both	Exterior Capsid	Structure
	Protease	Both	Core	Viral Maturation
L4	L4-100K	N/A	N/A	Hexon Assembly
	L4-33K	Empty	Unique Vertex	Packaging
	L4-22K	Empty	Unique Vertex	Packaging
	pVIII	Both	Interior Capsid	Cement
L5	pIV (Fiber)	Both	Exterior Vertex	Structure, Viral Entry

respectively. The other adenovirus late proteins are reviewed extensively by Ahi and Mittal (Ahi & Mittal, 2016).

1.1.6 Human Adenovirus E1A Structure

E1A is the first protein expressed after cell entry, and is tasked with changing the cellular landscape to become more conductive for viral replication (Nevins et al., 1979). This complex protein is organized into four regions: conserved region (CR) 1, CR2, CR3, and CR4. The E1A CR are segments of the protein that share sequence similarity between different HAdV types (Figure 1.3) (Avvakumov et al., 2002). These regions are largely intrinsically disordered except for CR3, which is predicted to form a Cys4 zinc finger, and the N-terminus of E1A from most HAdV types, which is predicted to form an α-helix (Figure 1.4) (Pelka et al., 2008). E1A is also alternatively spliced to form different protein products depending on the infection phase. In HAdV-5, there are 5 major isoforms of E1A: 289 residues (R), 243R, 217R, 171R, and 55R, which have mRNA sedimentation rates of 13S, 12S, 11S, 10S, and 9S respectively (Figure 1.4). 13S E1A, referred to as full length E1A for the remainder of this thesis, and 12S E1A are most prominent during the earlier phases of infection, while the remaining three isoforms are produced later in infection (Perricaudet et al., 1979; Stephens & Harlow, 1987).

1.1.7 Human Adenovirus E1A Transactivation of Viral and Host Genes

Being the first protein expressed, E1A is a strong transcriptional activator that starts the infection cycle by driving the expression of other early adenovirus genes (Berk et al., 1979; Jones & Shenk, 1979). E1A itself is controlled by a constitutively active enhancer (Hearing & Shenk, 1983). The regions responsible for transactivation have been mapped to the N-terminus and CR3 of E1A (Lillie & Green, 1989; Martin et al., 1990). E1A does not directly bind DNA and has no enzymatic activity; therefore, E1A relies on protein-protein interactions to mediate its effects (Ferguson et al., 1985). E1A targets various host

Figure 1.3. Sequence alignment of full length E1A from HAdV-3, 4, 5, 9, 12, 40, and 52. E1A sequence from representative adenovirus types from species A through G were aligned using Clustal Omega. The PXLXP motifs are outlined in red boxes, and the conserved regions are labelled below. Residues with high conservation are shown in blue, with darker shading indicating higher levels of conservation.

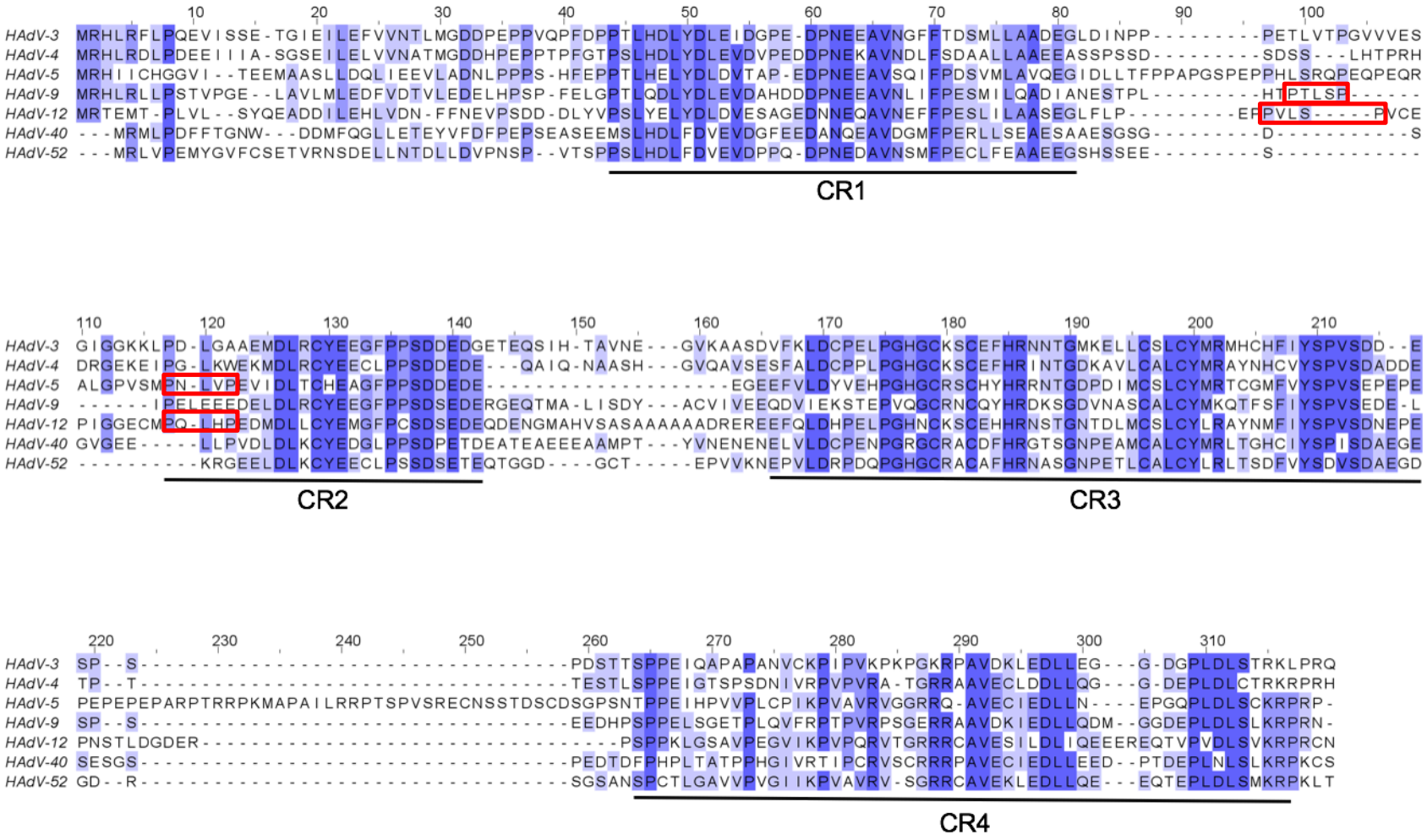
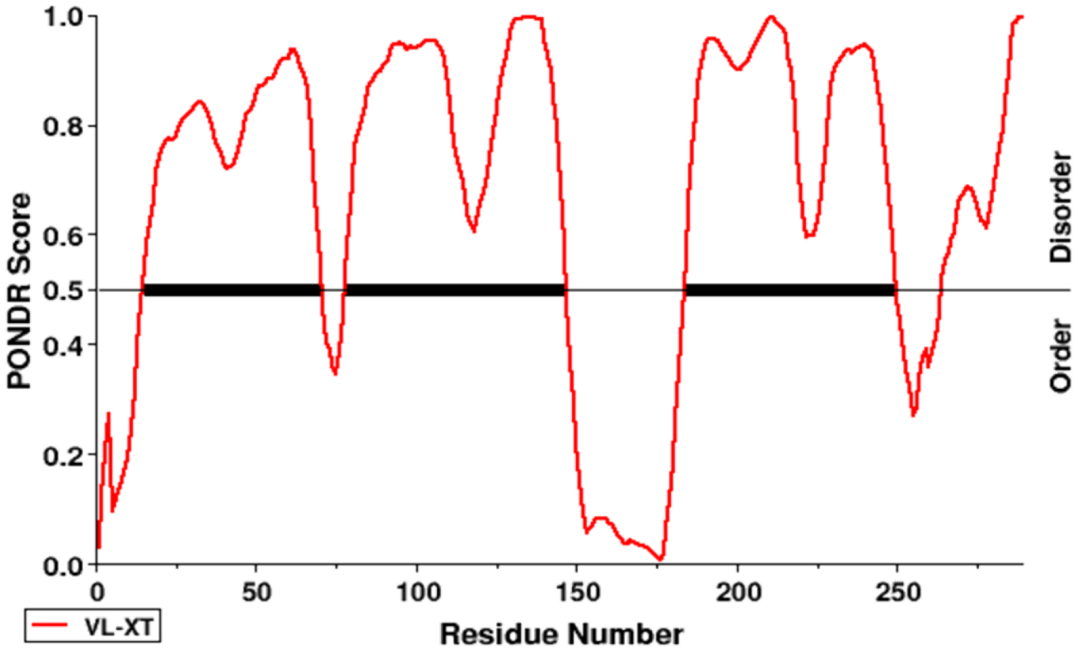
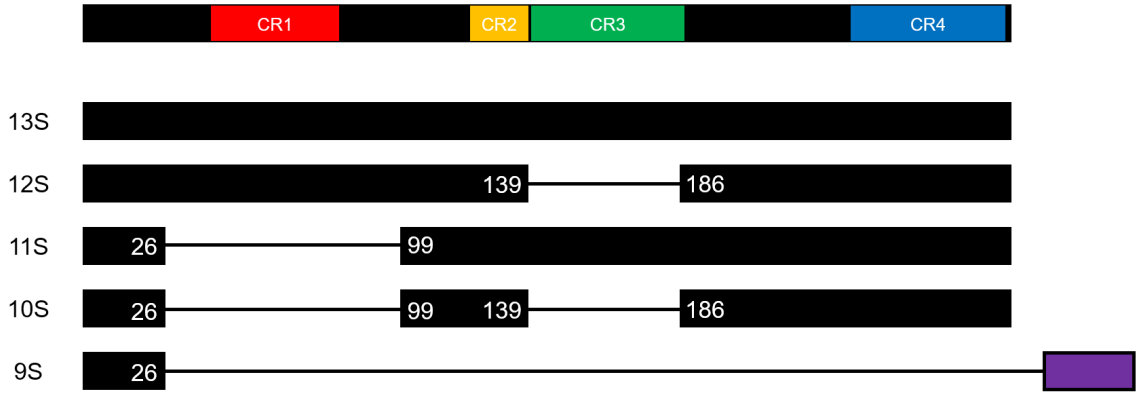

HAdV-3
HAdV-4
HAdV-5
HAdV-9
HAdV-12
HAdV-40
HAdV-52
CR1
CR2
CR3
CR4

Figure 1.4. PONDR graph of the five different E1A isoforms. Predictor of natural disordered regions (PONDR) is an online tool used to predict regions of order or disorder from the protein primary sequence. VL-XT is a combination of the VL1 predictor and the XN and XC predictors (Li et al., 1999; Romero et al., 1997). The PONDR score is between 1 and 0, with 1 being the ideal prediction of order and 0 being the idea prediction of disorder. The threshold is assigned at 0.5, where a protein is predicted to be disordered in the region with a score higher than 0.5. Here, HAdV-5 E1A primary sequence was the input. The four conserved regions are labelled in the bar below the PONDR graph. In HAdV-5 13S E1A, CR1 spans residues 40-80 (red), CR2 spans residues 121-139 (yellow), CR3 spans residues 140-188 (green), and CR4 spans residues 240-288 (blue). The five protein isoforms of E1A are also shown, with the residue position in reference to 13S E1A labelled in white where the alternative splicing of mRNA occurs. 9S E1A uses a different reading frame to code for the C-terminal region (shown in purple), resulting in a different protein primary sequence compared to the other E1A isoforms.


1.0
0.8
0.6
0.5
0.4
0.2
0.0
PONDR Score
Disorder
Order
0
50
100
150
200
250
Residue Number
VL-XT


CR1
CR2
CR3
CR4
13S
12S
139
186
11S
26
99
10S
26
99
139
186
9S
26

and viral promoters through an intrinsically-disordered promotor targeting subdomain, comprising of residues 183-188 in HAdV-5 E1A (Liu & Green, 1994; Webster & Ricciardi, 1991). This subdomain can be used to bind with various sequence-specific DNA-binding transcription factors such as members of the activating transcription factor (ATF) family, c-Jun, Sp1, and upstream stimulatory factor (USF) (Liu & Green, 1994). Subsequently, E1A stimulates the formation of the pre-initiation complex by recruiting the MED23 component of the mediator complex and TATA box binding protein (TBP) to the promotor region of target genes (Boyer et al., 1999; Cantin et al., 2003; Geisberg et al., 1994; Wang & Berk, 2002). The E1A N-terminus plays a similar role in facilitating transactivation by binding to various proteins that control transcription including activating protein 2 (AP2), thyroid hormone receptor (TR), p400, GCN5, TBP, and p300/CREB binding protein (Eckner et al., 1994; Fuchs et al., 2001; Lang & Hearing, 2003; Lipinski et al., 1998; Meng et al., 2003; Somasundaram et al., 1996).

1.1.8 Human Adenovirus E1A is a Viral Hub Protein

E1A directly binds to over 30 host factors to regulate cell cycle control, protein localization, and gene expression through disruption of various cell signaling networks (Figure 1.5) (Pelka et al., 2008). This capacity to bind with so many other proteins is especially impressive given that HAdV-5 E1A is less than 300 amino acids. As such, E1A is a hub protein, characterized by its ability to form a large number of interactions through short linear interaction motifs (SLiMs) (Nevins et al., 1979). SLiMs are typically 3 to 10 residues in length and are used by pathogens, including other viruses, to disrupt host cell signaling by mediating various protein-protein interactions (Davey et al., 2011; Via et al., 2015). These motifs are typically present on intrinsically disordered regions of a protein, but may also be found on solvent-exposed surfaces of alpha helices (Van Roey et al., 2014). Due to their short length, SLiMs typically mediate interactions that are transient or of low affinity; therefore, some protein-protein interactions mechanisms require more than one SLiM to facilitate stable binding (Van Roey et al., 2014). Due to the simplicity of many of these motifs, SLiMs are an avenue for molecular mimicry and are conducive to evolutionary plasticity (Neduva & Russell, 2005). Pathogens, especially viruses due to

Figure 1.5. E1A protein interaction network. Graphical representation of the E1A primary and secondary interactors. Graph was created using Gephi 0.9.1 with data from BioGRID build 3.4.144. E1A is shown in the white circle in the center of the diagram. Primary interactors are represented as larger circles near the outer region of the diagram, and are supported by at least two peer-reviewed publications. The relative sizes of the primary interactors are proportional to its number of binding partners. The primary interactors are sorted into 11 groups based on sub-networks using Gephi. The secondary interactors are represented by smaller circles scattered throughout the diagram and are coloured and positioned closely to the E1A primary interactor with which it binds. There are 31 primary interactors and 2125 unique secondary interactors represented in this diagram. Graph was created by Dr. Joe Mymryk (University of Western Ontario, unpublished, 2017).

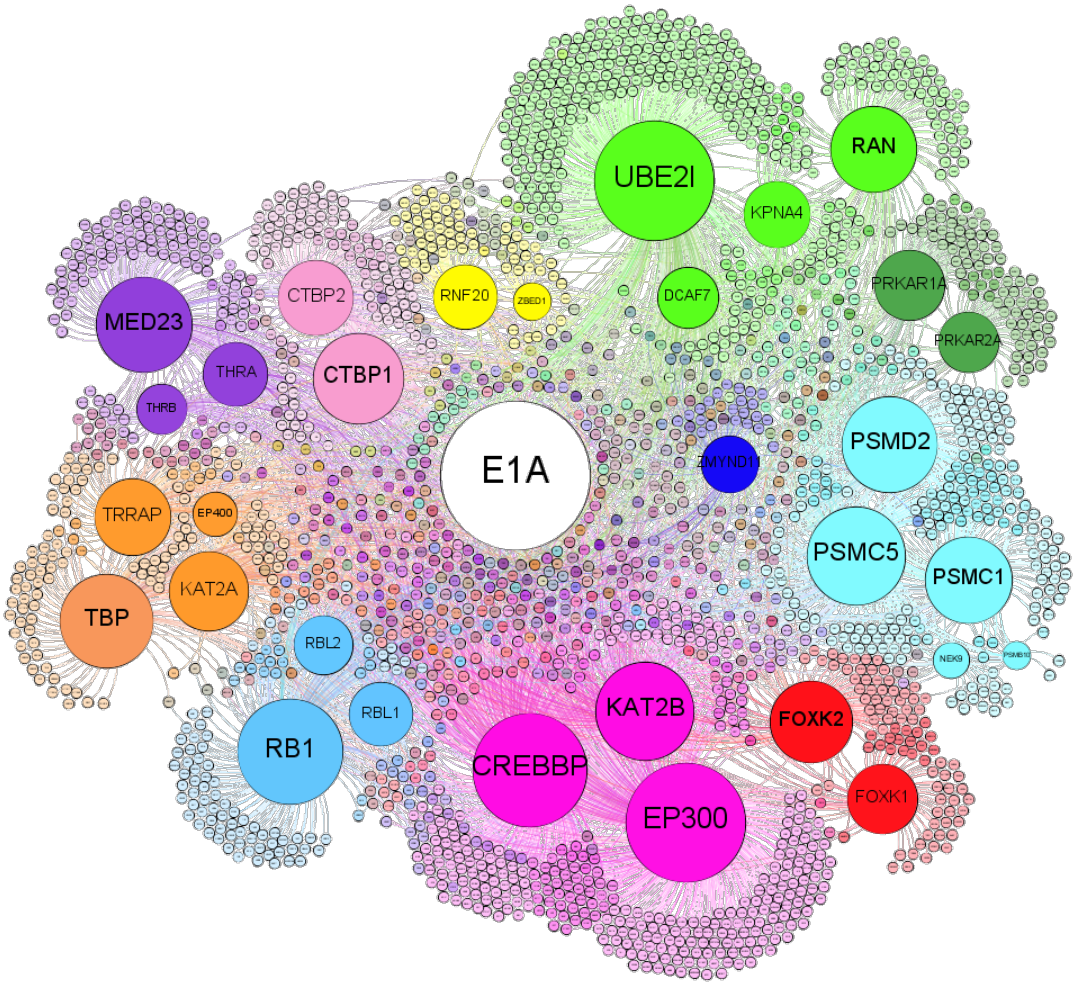
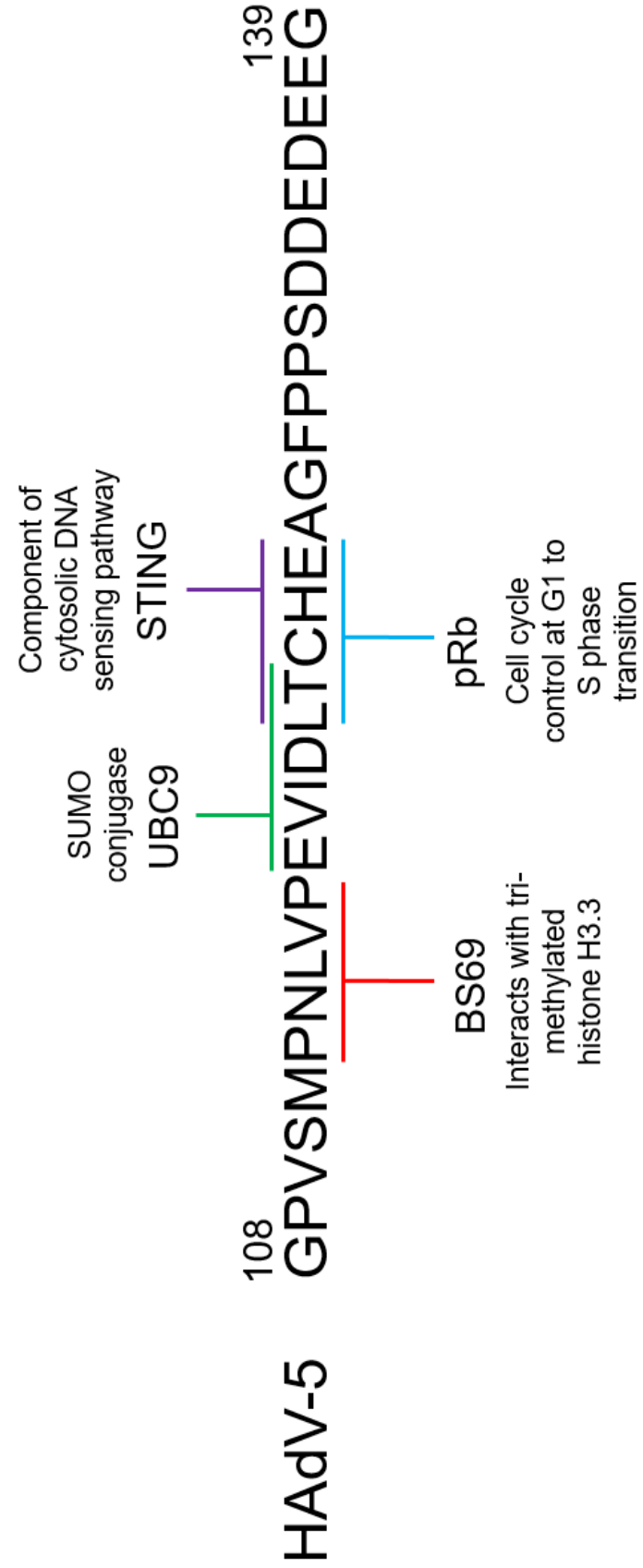

UBE2I
RAN
KPNA4
MED23
CTBP2
RNF20
ZBED1
DCAF7
PRKAR1A
PRKAR2A
THRA
CTBP1
THRB
E1A
ZMYND11
PSMD2
TRRAP
EP400
PSMC5
PSMC1
KAT2A
TBP
RBL2
NEK9
PSMB10
KAT2B
FOXK2
RBL1
RB1
CREBBP
FOXK1
EP300

Figure 1.6. Diagram of E1A interaction partners and short linear motifs within conserved region 2. Protein sequence of HAdV-5 E1A (residues 108 – 139) is shown. The PXLXP, EVIDLT, and LXCXE motifs are matched with their respective binding partners along with a brief description of the function of BS69, UBC9, pRB, and STING.


HAdV-5
108
GPVSMPNLVPEVIDLTCHEAGFPPSDDEDEEG
139
SUMO conjugase
UBC9
Component of cytosolic DNA sensing pathway
STING
pRb
Cell cycle control at G1 to S phase transition
BS69
Interacts with tri-methylated histone H3.3

their small and rapidly-evolving genomes, frequently exploit these motifs to hijack specific regulatory pathways within the host cell to create a favourable environment for pathogen survival and propagation (Neduva & Russell, 2005). SLiMs are particularly common in HAdV-5 E1A CR2, as the PXLXP, EVIDLT, and LXCXE motifs exist here in tandem (Figure 1.6). The PXLXP and EVIDLT motifs are used to interact with BS69 and small ubiquitin-like moiety (SUMO) conjugase ubiquitin conjugase 9 (UBC9) respectively, while the LXCXE motif is used to interact with both pRb and stimulator of interferon genes (STING) (Ansieau & Leutz, 2002; Avvakumov et al., 2004; Lau et al., 2015; Yousef et al., 2010).

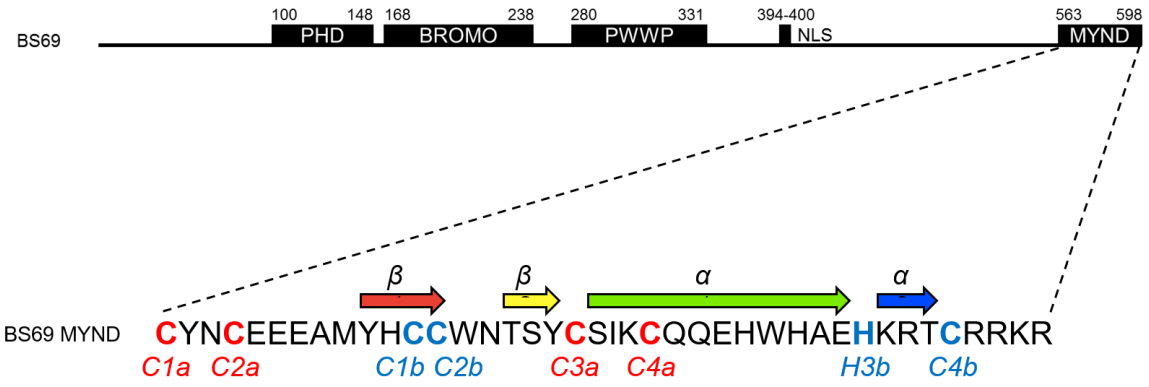
1.2 Human BS69

1.2.1 Discovery and Structure of BS69

BS69, also known as ZMYND11, was initially identified in 1995 as an E1A interacting protein (Hateboer et al., 1995). This protein binds directly to E1A CR2 using the PXLXP motif, but there is conflicting data in the literature regarding the involvement of CR3 in this interaction (Ansieau & Leutz, 2002; Hateboer et al., 1995). BS69 was first described as a strong inhibitor of E1A transactivation; however, it was later found that the repressive function of BS69 is not limited to just E1A mediated transcriptional activation (Hateboer et al., 1995; Masselink & Bernards, 2000). Recently, BS69 has also been found to be involved in epigenetic gene regulation (Guo et al., 2014; Wen et al., 2014). BS69 binds specifically to histone H3.3 trimethylated at lysine residue 36 (H3.3K36me3) to cause localized gene repression (Guo et al., 2014; Wen et al., 2014).

BS69 has a molecular weight of 69kDa, and the full-length protein is 602 residues in length. BS69 is localized to the nucleus, is ubiquitously expressed, and carries out a variety of functions associated with gene regulation (Velasco et al., 2006). This protein consists of 4 previously described domains: plant homeodomain (PHD); bromodomain (BROMO); PWWP domain; and Myeloid, Nervy, and deformed epithelial autoregulatory factor 1 homolog (DEAF-1) (MYND) domain (Figure 1.7, 1.8). There is also a zinc finger between the BROMO and PWWP domains, and a coiled-coil structure and a nuclear localization

Figure 1.7. Structure of BS69 MYND domain. Diagram of the BS69 domains and nuclear localization sequence (NLS) are shown with the residue positions above. The sequence of the MYND domain is below, with the cysteine and histidine residues used to coordinate one of the two zinc ions bolded, labelled, and coloured based on which zinc ion they coordinate. The MYND domain is composed of two antiparallel beta sheets (red and yellow) followed by two alpha helices (green and blue). The NMR structure (PDB: 2HDA) shows the secondary structures colour coordinated with the sequence above, along with the two zinc ions are grey spheres, and the residues used to coordinate these ions. Adapted from (Harter et al., 2016).


BS69
100
148
168
238
280
331
394-400
563
598
PHD
BROMO
PWWP
NLS
MYND
β
β
α
α
BS69 MYND
CYNCEEEAMYHCCWNTSYCSIKCQQEHWHAEHKRTCRRKR
C1a C2a C1b C2b C3a C4a H3b C4b

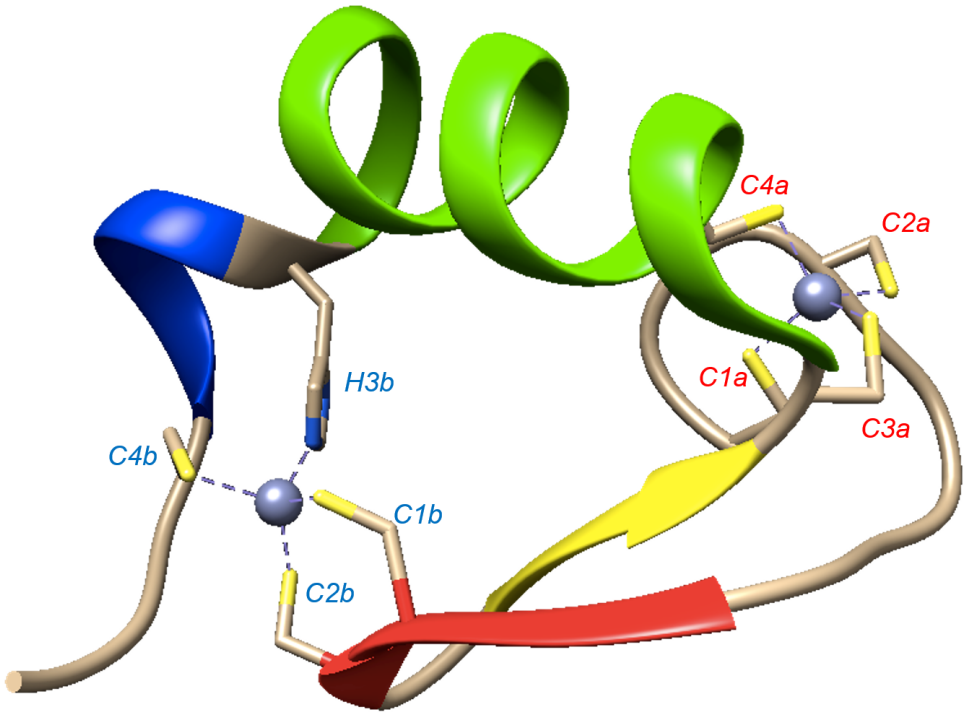
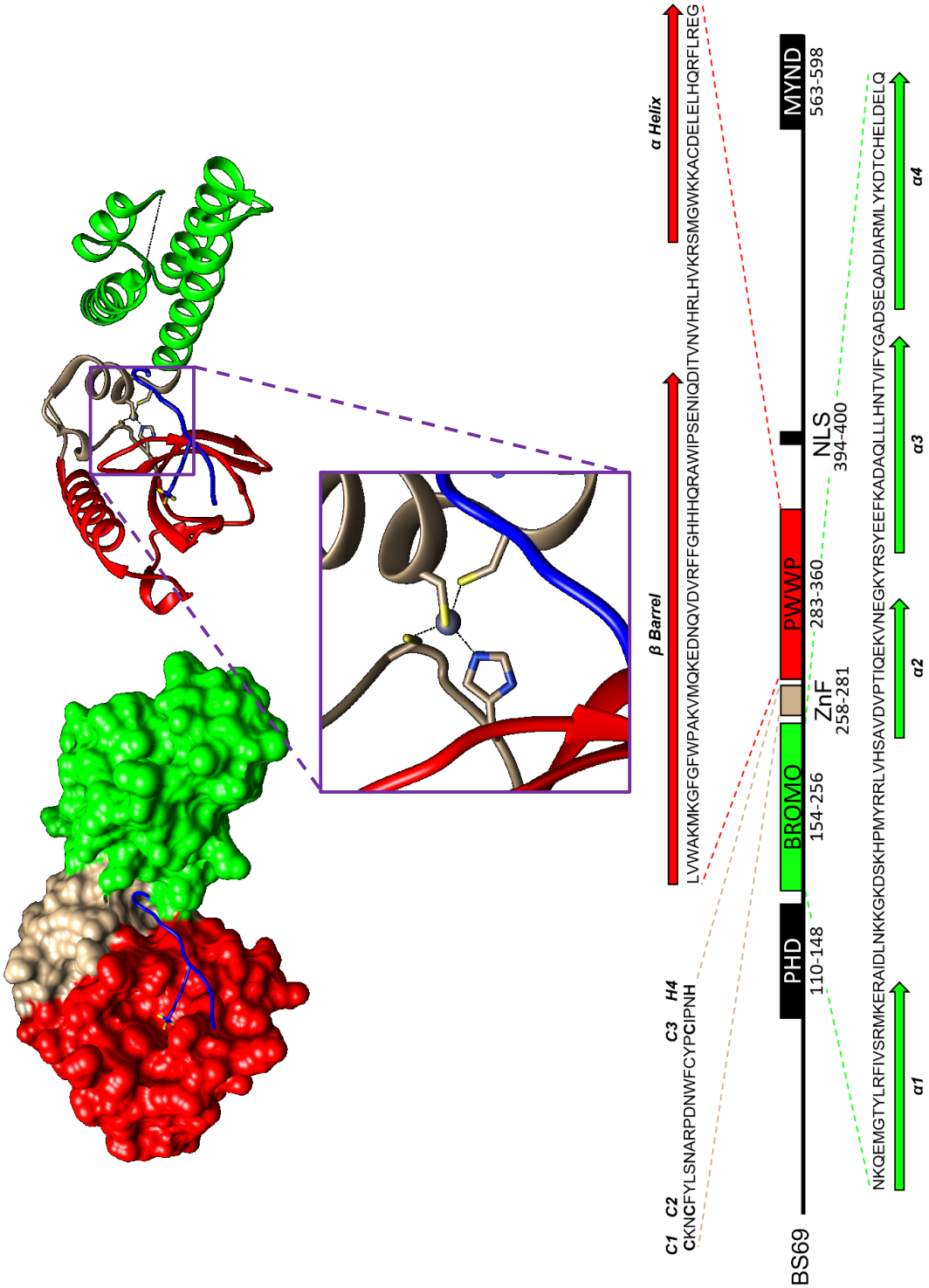

C4a
C2a
H3b
C1a
C3a
C4b
C1b
C2b

Figure 1.8. Structure of the BROMO and PWWP domains of BS69. NMR structure (PDB: 4N4I) of the BS69 BROMO (green) and PWWP (red) domains in both globular and ribbon models (Wen et al., 2014). The zinc finger (ZnF) between the domains is shown in beige, with an expanded view to show the cysteine and histidine residues used to coordinate the zinc ion. A histone H3.3K36me3 peptide is shown in blue. The BROMO domain is composed of four alpha helices, while the PWWP domain consists of a beta barrel followed by an alpha helix. The bottom diagram on the bottom shows the sequence of this region in BS69, along with the location of the secondary structures. The residue position is shown below the respective domains. NLS = nuclear localization sequence. Adapted from (Wen et al., 2014).


α Helix
β Barrel
LVWAKMKGFGFWPAKVMQKEDNQVDVRFFGHHHQRAWIPSENIQDITVNVHRLHVKRSMGWKKACDELELHQRFLREG
C1 C2
CKNCFYLSNARPDNWFCYPCIPNH
C3 H4
BS69
PHD
110-148
BROMO
154-256
ZnF
258-281
PWWP
283-360
NLS
394-400
MYND
563-598
α1
α2
α3
α4
NKQEMGTYLRFIVSRMKERAIDLNKKGKDSKHPMYRRLVHSAVDVPTIQEKVNEGKYRSYEEFKADAQLLLHNTVIFYGADSEQADIARMLYKDTCHELDELQ

signal (NLS) between the PWWP and MYND domains (Harter et al., 2016; Wang et al., 2014). This protein is highly conserved between humans and mice, especially the MYND domains, which share 100% sequence identity (Figure 1.9). The BS69 splice variant, bone morphogenetic protein (BMP) receptor associated molecule 1 (BRAM1), consists of 12 unique N-terminal residues and 185 shared C-terminal residues with BS69 (Kurozumi et al., 1998). BRAM1 localizes to the cytoplasm and interacts with BMP type IA receptor and plays a role in the BMP signaling pathway (Kurozumi et al., 1998). This pathway is required for normal functioning of bone tissues, as its perturbation is correlated with various bone diseases (Sánchez-Duffhues et al., 2015).

1.2.2 The Domains and Functions of BS69

The MYND domain of BS69 is 40 residues in length, and contains two zinc ions coordinated by seven cysteine residues and one histidine residue (Figure 1.7). This domain contains two anti-parallel beta sheets followed by two alpha helices. The MYND domain of BS69 primarily functions to mediate protein-protein interactions with binding partners such as E1A, Epstein Barr Nuclear Antigen 2 (EBNA2), nuclear receptor co-repressor 1 (N-CoR), ETS2, B-myb, C-myb, and zinc fingers and homeoboxes (ZHX) 1 (Ansieau & Leutz, 2002; Harter et al., 2016; Hateboer et al., 1995; Ladendorff et al., Lipsick, 2001; Masselink & Bernards, 2000; Masselink et al, 2001; Ogata-Kawata et al., 2007; Wei et al., 2003). E1A, EBNA2, and ZHX1 all use the PXLXP motif to interact with the MYND domain; additionally, it is interesting to note that N-CoR and ETS2 also contain PXLXP motifs in their primary sequence. E1A and EBNA2 both facilitate transcriptional activation, and the interaction with the MYND domain of BS69 represses this function (Harter et al., 2016; Hateboer et al., 1995). Similarly, BS69 also inhibits transcriptional activation of B-myb and C-myb (Ladendorff et al., 2001; Masselink et al., 2001). The mechanism for this repression is not well understood, but may involve the recruitment of the N-CoR repressor (Masselink & Bernards, 2000). The opposite is also true, where expression of E1A and EBNA2 relieves BS69-mediated gene repression (Ansieau & Leutz, 2002; Masselink & Bernards, 2000). Interestingly, through this interaction, BS69 has also

Figure 1.9. Differences in protein sequence between human BS69 and murine BS69. Black rectangles represent the various BS69 domains and the NLS, with numbers on the corners representing the residue positions. Sequence changes below denote changes in protein sequence from human BS69 to murine BS69.


BS69
100
PHD
148
168
BROMO
238
280
PWWP
331
394-400
NLS
563
MYND
598
N98T
P139H
N152H
T153S
N183S
I332V
K433R
N459S
N519S
M520L
E523D
C528G
V531L
T548A

been found to stabilize E1A by inhibiting ubiquitin-dependent degradation of E1A (Isobe et al., 2006).

The activity of several other cellular transcription factors is influenced by the versatile BS69 protein. ETS1 and ETS2 are signal-dependent transcription factors that promote expression of genes downstream of the rat sarcoma (Ras) / mitogen-activated protein kinase (MAPK) signaling pathway (Plotnik et al., 2014). Both these transcription factors have the PXLXP motif near the N-terminus of their primary sequences, but only ETS2 binds to BS69 (Plotnik & Hollenhorst, 2017; Wei et al., 2003). As a consequence, BS69 preferentially inhibits ETS2-mediated transactivation, making ETS2 a transcriptional repressor or weaker activator compared to ETS1 (Plotnik & Hollenhorst, 2017; Wei et al., 2003). Like the interaction pattern seen with ETS2 and ETS1, BS69 is able to bind to ZHX1 but not ZHX2 despite both proteins having a PXLXP motif (Kawata et al., 2003; Ogata-Kawata et al., 2007). Thus, the MYND domain of BS69 binds to various proteins using the core PXLXP motif, but it is evident that the context of the motif is also important in mediating a strong interaction. The C-terminus of BS69 also binds to several chromatin remodeling factors such as BRG1, EZH2, and HDAC1; therefore, BS69 may also play a role in chromatin remodeling (Guo et al., 2014; Velasco et al., 2006).

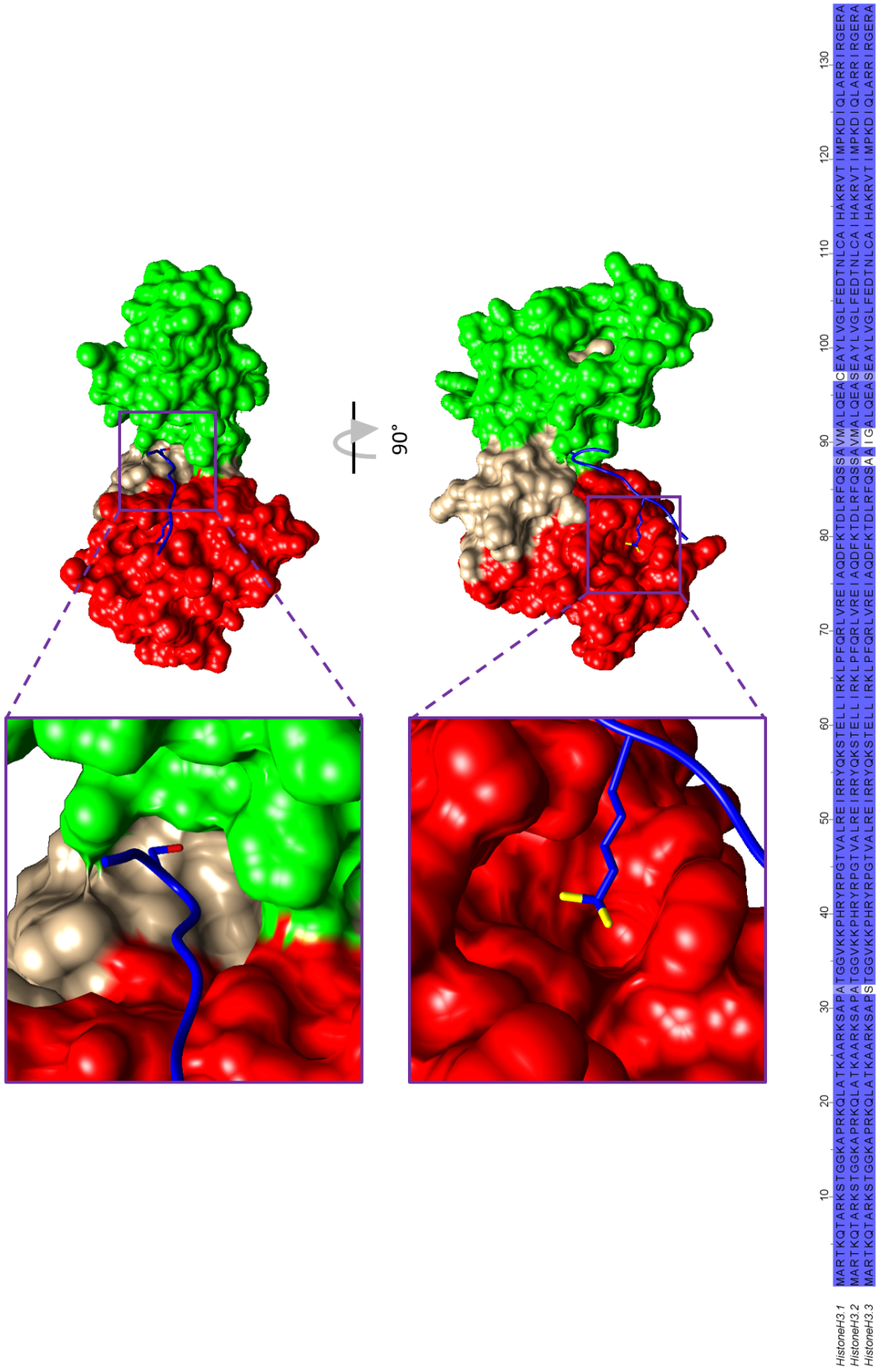
The PHD, BROMO, and PWWP domains are located in tandem near the N-terminus of the protein (Figure 1.8). These domains are also used to facilitate protein-protein interactions, most notably with histone tails (Guo et al., 2014; Wen et al., 2014). The PHD domain was first discovered in 1993 in *Arabidopsis thaliana*; since then, more than 170 proteins with these domains have been identified in the human genome (Liu et al., 2012; Schindler et al., 1993). PHD fingers are primarily involved in recognizing various acetylated and di- or trimethylated histone tails (Bortoluzzi et al., 2017; Li & Li, 2012). Proteins containing a BROMO domain are also prevalent, as there are currently 61 identified modules in 42 proteins in the human genome (Fujisawa & Filippakopoulos, 2017). These domains are also involved in epigenetic regulation of gene expression, although they selectively recognize acetylated lysine residues on histone tails (Fujisawa & Filippakopoulos, 2017).

Lastly, the PWWP domain is named after a conserved Pro-Trp-Trp-Pro motif found in this domain. This domain was first discovered in the histone methyltransferase NSD2 in 1998 (Stec et al., 1998). Since then, the PWWP domain has been found in over 20 proteins in the human genome (Qin & Min, 2014). This domain is capable of binding methylated lysine residues on histone tails, and interacting directly with DNA (Qin & Min, 2014). In the BS69 protein, these three domains fold together to form a functional module, as they are all involved in the interaction with H3.3K36me3 (Guo et al., 2014; Wang et al., 2014; Wen et al., 2014). From two crystal structures published by two groups, it is evident that the PWWP domain is primarily responsible for the binding specificity between BS69 and H3.3K36me3 (Figure 1.10) (Wang et al., 2014; Wen et al., 2014). There are five differences in the primary sequences between histone H3.3 and histone H3.1/H3.2, and only one difference is within the tail region. Histone H3.3 contains a serine residue at position 32, while histone H3.1/H3.2 contain alanine residues at this position (Figure 1.10). The serine residue in histone H3.3 binds to the interface between the zinc finger and the PWWP domain of BS69, and is crucial in the formation of a strong interaction between these two proteins; therefore, post-translational modification to this serine residue, such as phosphorylation, may play a role in the regulation of this protein-protein interaction (Guo et al., 2014; Wen et al., 2014).

1.2.3 BS69 Uses an Epigenetic Mechanism to Downregulate Gene Expression

BS69 downregulates gene expression of genes decorated with H3.3K36me3 through two mechanisms: suppression of RNA polymerase elongation and upregulation of intron retention (Guo et al., 2014; Wen et al., 2014). Using RNA-seq analysis, Wen *et al*. noted that RNAPII elongation in gene bodies was greater in BS69-depleted cells compared to a wildtype control, while the density of RNAPII in respective promoter regions were unchanged (Wen et al., 2014). These BS69-depleted cells also have increased elongation specific RNAPII levels, denoted by phosphorylation of serine residue at position 2 in the unique heptameric YSPTSPS sequence in the C-terminal domain of RNAPII (Ahn et al., 2004; Wen et al., 2014). Lastly, chromatin immunoprecipitation sequencing (ChIP-seq)

Figure 1.10. Interaction between the BROMO and PWWP domains of BS69 with H3.3K36me3. NMR structure (PDB: 4N4I) of the BS69 BROMO (green) and PWWP (red) domains (Wen et al., 2014). The zinc finger (ZnF) between the domains is shown in beige. A histone H3.3K36me3 peptide is shown in blue. The S32 residue is shown zoomed in on top, while the bottom panel shows trimethylated K36 residue. Sequence alignment of Histone H3.1, H3.2, and H3.3 is shown on the bottom. Residues with high conservation are shown in blue, with darker shading indicating higher levels of conservation. Adapted from (Wen et al., 2014).


90°
HistoneH3.1
HistoneH3.2
HistoneH3.3
MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA
MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA
MARTKQTARKSTGGKAPRKQLATKAARKSASSTGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAIGALQEASEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA
10
20
30
40
50
60
70
80
90
100
110
120
130

showed that BS69 co-localizes with histone H3.3K36me3 on transcriptionally active genes (Guo et al., 2014; Wen et al., 2014). Altogether, these results demonstrate that BS69 negatively regulates RNAPII elongation in genes decorated with histone H3.3K36me3 (Wen et al., 2014).

RNA splicing is also regulated by BS69 in genes decorated with the H3.3K36me3 marker (Guo et al., 2014). Using Co-immunoprecipitation (Co-IP), Guo *et al*. showed that chromatin-bound BS69 interacts with various components of the spliceosome, which is a molecular complex used to mediate RNA splicing, including elongation factor Tu GTP-binding domain-containing protein 2 (EFTUD2), pre-mRNA processing splicing factor 8 (PRPF8), U5 small nuclear ribonucleoprotein 200 kDa helicase (SNRNP200), serine/arginine-rich splicing factor 1 (SRSF1), and pinin (PNN) (Guo et al., 2014). EFTUD2, PRPF8, and SNRNP200 are all components of the U5 spliceosome, while SRSF1 and PNN are members of a conserved family of serine and arginine-rich proteins also involved in RNA splicing (Guo et al., 2014). Using RNA-seq, the group found that BS69-depleted cells exhibited decreased intron retention events compared to wildtype cells, and that this observation is dependent on the interaction between BS69 and the U5 spliceosome component, EFTUD2 (Guo et al., 2014). Furthermore, BS69 is enriched at these alternatively spliced genes, and intron retention of these transcripts causes nonsense-mediated mRNA decay (Guo et al., 2014). In summary, these studies have demonstrated two novel mechanisms by which BS69 is able use epigenetic mechanisms to regulate gene expression.

1.2.4 Disease Association of BS69 and Histone H3.3 Mutants

Mutations in either BS69 or histone H3.3 have also been linked to various diseases. Mutations in histone H3.3 are correlated with pediatric brain cancers and bone cancers; specifically, K27M, G34R/V/W/L, and K36M mutations in the histone tail of histone H3.3 is associated with glioblastoma multiforme, chondroblastoma, giant-cell tumour of the bone, and human papillomavirus-negative head and neck squamous cell carcinoma (Behjati et al., 2013; Fontebasso et al., 2014; Papillon-Cavanagh et al., 2017;

Schwartzentruber et al., 2012). Interestingly, the K27M mutation decreases methylation of the lysine residue 36 in histone H3.3 tails in human embryonic kidney 293 (HEK293) cells as well (Chan et al., 2013). Additionally, a consequence of the K36M mutation prevalent in giant-cell tumour of the bone is that BS69 can no longer bind to the histone tail (Behjati et al., 2013; Wang et al., 2014). Mutations and copy number variations in BS69 have also been linked to cognitive / developmental delay and various hematological malignancies; however, a mechanistic connection between abrogation of BS69 function and disease has not yet been made (Cobben et al., 2014; Coe et al., 2014; DeScipio et al., 2012; Moskowitz et al., 2016; Walter et al., 2009; Yang et al., 2010).

1.3 Rationale, Hypothesis, and Objectives

BS69 was first identified in 1995 as an E1A interactor through a pulldown screen performed by the Bernards group (Hateboer et al., 1995). The protein was described as an inhibitor of transactivation, but the mechanism by which this occurs and the genetic targets are unclear. Furthermore, this interaction has also only been studied in the prototypical HAdV-5 of species C, and it is uncertain if this interaction is conserved in other adenovirus species. The binding surface is also poorly described, making it difficult to construct the appropriate E1A mutants required to study this interaction. In 2014, two publications showed that BS69 is also a histone reader that specifically targets histone H3.3 trimethylated at lysine residue 36 (H3.3K36me3) (Guo et al., 2014; Wen et al., 2014).Therefore, this E1A binding protein may also be involved in epigenetic mechanisms of gene regulation. Thus, I set out to characterize this interaction and examine its effects on E1A mediated gene transcription.

Based on the rationale, I hypothesize that the interaction between E1A and BS69 alters transcriptional regulation of viral and host genes. To investigate my hypothesis, I have formed three objectives:

1. Determine the specific residues of E1A required to bind with BS69 and the conservation of this interaction in different human adenovirus species.
2. Examine the effects of BS69 on E1A mediated transcriptional activation.
3. Construct a BS69-deficient cell line to study changes in BS69-mediated gene expression.

2 Materials and Methods

2.1 Generation of Plasmids

Plasmids 8-15 (Table 2.1) were generated by combining self-annealing oligonucleotides and ligating them into the pBAIT (plasmid 1, Table 2.1) backbone vector (Zhang et al., 2001). Oligonucleotides (Table 2.2) were first diluted to 10 µM and were phosphorylated using 10U of polynucleotide kinase (PNK) (New England Biolabs) in T4 DNA ligase buffer (New England Biolabs) in a 20µl reaction at 37^{o}C for 1 hour. The kinase was then deactivated by incubating the mixture at 65^{o}C for 20 minutes. The phosphorylated self-annealing oligonucleotide pairs (Table 2.2) were mixed together and incubated at a final concentration of 0.5µM each primer in 100µl of ddH_2O at 98^{o}C for 2 minutes, then cooled gradually to 25^{o}C over 45 minutes using a thermocycler (SimpliAmp from Applied Biosystems). 2µg of the pBAIT backbone vector was then digested using EcoRI-high fidelity (HF) and SalI-HF (New England Biolabs) at 37^{o}C for 1 hour in CutSmart buffer (New England Biolabs). After digestion, the DNA was resolved on an agarose gel and the bands of interest were cut and processed using EZ-10 Spin Columns (BioBasic) according to the manufacturer's protocol. The double-stranded oligonucleotides were then diluted 1:100 in ddH_2O before being ligated into the corresponding sites of pBAIT using the Quick Ligation Kit (New England Biolabs) according to the manufacturer's protocol.

Plasmids (16-30) were generated using overlap extension (Ho et al., 1989). Polymerase chain reaction (PCR) amplification of E1A mutants were completed using Phusion Polymerase (Thermo Fisher) according to the manufacturer's protocol. The first round of PCR was completed using 10µM each of the matching forward and reverse primers (Table 2.3), and 250ng template DNA in a 50µl reaction. The reaction underwent 35 cycles with the annealing and extension steps completed at 60^{o}C for 30 seconds and 72^{o}C for 15 seconds respectively. The PCR mix was then incubated with 1µl DpnI (New England Biolabs) at 37^{o}C for 1 hour to digest the plasmid template before the PCR products were purified using the GeneJET PCR Purification Kit (Thermo Fisher) according to the manufacturer's protocol. Matching pairs

Table 2.1. List of Plasmids Used in this Study

#	Name	Backbone	Description / Insert
1	pBAIT	pBAIT	Bait vector used for yeast two-hybrid assays. Contains N-terminal LexA DBD to the MCS, AmpR, LEU2.
2	pBAIT HAdV-3 E1A CR2	pBAIT	aa 85-150 of HAdV-3 E1A
3	pBAIT HAdV-4 E1A CR2	pBAIT	aa 91-145 of HAdV-4 E1A
4	pBAIT HAdV-5 E1A CR2	pBAIT	aa 93-139 of HAdV-5 E1A
5	pBAIT HAdV-9 E1A CR2	pBAIT	aa 84-138 of HAdV-9 E1A
6	pBAIT HAdV-12 E1A CR2	pBAIT	aa 84-146 of HAdV-12 E1A
7	pBAIT HAdV-40 E1A CR2	pBAIT	aa 78-142 of HAdV-40 E1A
8	pBAIT E1A T1	pBAIT	aa 111-120 of HAdV-5 E1A
9	pBAIT E1A T2	pBAIT	aa 112-120 of HAdV-5 E1A
10	pBAIT E1A T3	pBAIT	aa 111-119 of HAdV-5 E1A
11	pBAIT E1A T4	pBAIT	aa 112-119 of HAdV-5 E1A
12	pBAIT E1A T5	pBAIT	aa 112-118 of HAdV-5 E1A
13	pBAIT E1A T6	pBAIT	aa 113-118 of HAdV-5 E1A
14	pBAIT E1A T7	pBAIT	aa 112-117 of HAdV-5 E1A
15	pBAIT E1A T8	pBAIT	aa 113-117 of HAdV-5 E1A
16	pBAIT E1A M1	pBAIT	S111A CR2 of HAdV-5 E1A
17	pBAIT E1A M2	pBAIT	M112A CR2 of HAdV-5 E1A
18	pBAIT E1A M3	pBAIT	P113A CR2 of HAdV-5 E1A
19	pBAIT E1A M4	pBAIT	N114A CR2 of HAdV-5 E1A
20	pBAIT E1A M5	pBAIT	L115A CR2 of HAdV-5 E1A
21	pBAIT E1A M6	pBAIT	V116A CR2 of HAdV-5 E1A

Table 2.1. List of Plasmids Used in this Study (Continued)

22	pBAIT E1A M7	pBAIT	P117A CR2 of HAdV-5 E1A
23	pBAIT E1A M8	pBAIT	E118A CR2 of HAdV-5 E1A
24	pBAIT E1A M9	pBAIT	V119A CR2 of HAdV-5 E1A
25	pBAIT E1A M10	pBAIT	I120A CR2 of HAdV-5 E1A
26	pBAIT E1A M11	pBAIT	P113A L115A P117A CR2 of HAdV-5 E1A
27	pBAIT E1A M12	pBAIT	M112I L115A CR2 of HAdV-5 E1A
28	pBAIT E1A M13	pBAIT	P113A L115A CR2 of HAdV-5 E1A
29	pBAIT E1A M14	pBAIT	M112I L115A CR2 of HAdV-5 E1A
30	pBAIT E1A M15	pBAIT	L115A P117A CR2 of HAdV-5 E1A
31	pJG4-5+	pJG4-5+	Prey vector used for yeast two-hybrid assays. Contains N-terminal B42 activation domain to MCS, HA Tag, AmpR, TRP1. GAL1 promoter.
32	pJG4-5+ BS69 MYND	pJG4-5+	BS69 MYND domain aa 427-602
33	pSH18-34	pSH18-34	Yeast two-hybrid reporter plasmid. Contains 8x LexA binding sites upstream
34	pEGFP	pEGFP	Mammalian plasmid to express N-terminal EGFP to the MCS, KanR.
35	pEGFP 13S E1A	pEGFP	Full length HAdV-5 E1A
36	pEGFP CR2 E1A	pEGFP	aa 93-139 of HAdV-5 E1A
37	pEGFP CR2-CR3 E1A	pEGFP	aa 93-204 of HAdV-5 E1A (insert generated by PCR from plasmid 35)
38	pcDNA3 HA	pcDNA3 HA	Mammalian plasmid to express N-terminal HA tag to the MCS, AmpR.
39	pcDNA3 HA BS69 Full Length	pcDNA3 HA	BS69 Full Length aa 46-602 (insert generated by PCR from plasmid 53)
40	pcDNA3 HA BS69 MYND	pcDNA3 HA	BS69 MYND domain aa 427-602

Table 2.1. List of Plasmids Used in this Study (Continued)

41	pcDNA3 HA BS69 No MYND	pcDNA3 HA	BS69 No MYND domain aa 46-426 (insert generated by PCR from plasmid 53).
42	pET28a BS69	pET28a	Protein purification plasmid. Contains N-terminal 6x His tag, thrombin cleavage site, BS69 MYND domain aa 427-602, KanR).
43	pGEX-4T1 BS69 Full Length	pGEX-4T1	Protein purification plasmid. Contains N-terminal GST tag, thrombin cleavage site, BS69 Full Length aa 46-602, AmpR.
44	pGEX-4T1 BS69 MYND	pGEX-4T1	BS69 MYND domain aa 427-602
45	pGEX-4T1 Rb	pGEX-4T1	Rb small pocket aa 385-773
46	pGEX-4T1 UBC9	pGEX-4T1	UBC9 full length aa 1-158
47	pM	pM	Mammalian plasmid to express N-terminal GAL4 DBD (aa 1-147) to the MCS, AmpR.
48	pGL2-(UAS)6-Luc	pGL2-(UAS)6-Luc	Mammalian luciferase reporter plasmid. Contains 6x GAL4 binding sites upstream of the luciferase reporter gene.
49	pM 13S Ad5 E1A	pM	Full length HAdV-5 E1A
50	pM 13S Ad12 E1A	pM	Full length HAdV-12 E1A
51	pM 13S Ad40 E1A	pM	Full length HAdV-40 E1A
52	pM 13S E1A L115A	pM	L115A full length HAdV-5 E1A
53	pRc/CMV BS69	pRc/CMV	Mammalian expression vector containing full length BS69, AmpR.
54	pSV-β-galactosidase	pSV	Internal control vector for luciferase assays. SV40 promoter and enhancer upstream of the lacZ gene. AmpR

Note: AmpR = ampicillin resistance gene, KanR = kanamycin resistance gene, MCS = multiple cloning site, UAS = upstream activation sequence, DBD = DNA binding domain

Table 2.2. List of Self-Annealing Oligonucleotides Used in this Study

Construct	Oligonucleotide Sequences (5' → 3')
E1A T1	AATTCTCTGGTGGTTCTATGCCAAACCTTGTACCGGAGGTGATCTAAG (F) TCGACTTAGATCACCTCCGGTACAAGGTTTGGCATAGAACCACCAGAG (R)
E1A T2	AATTCTCTGGTGGTATGCCAAATTTAGTTCCAGAAGTTATCTAAG (F) TCGACTTAGATAACTTCTGGAACTAAATTTGGCATACCACCAGAG (R)
E1A T3	AATTCTCTGGTGGTTCTATGCCAAATTTAGTTCCAGAAGTTTAAG (F) TCGACTTAAACTTCTGGAACTAAATTTGGCATAGAACCACCAGAG (R)
E1A T4	AATTCTCTGGTGGTATGCCAAATTTAGTTCCAGAAGTTTAAG (F) TCGACTTAAACTTCTGGAACTAAATTTGGCATACCACCAGAG (R)
E1A T5	AATTCTCTGGTGGTATGCCAAATTTAGTTCCAGAATAAG (F) TCGACTTATTCTGGAACTAAATTTGGCATACCACCAGAG (R)
E1A T6	AATTCTCTGGTGGTCCAAATTTAGTTCCAGAATAAG (F) TCGACTTATTCTGGAACTAAATTTGGACCACCAGAG (R)
E1A T7	AATTCTCTGGTGGTATGCCAAATTTAGTTCCATAAG (F) TCGACTTATGGAACTAAATTTGGCATACCACCAGAG (R)
E1A T8	AATTCTCTGGTGGTCCAAATTTAGTTCCATAAG (F) TCGACTTATGGAACTAAATTTGGACCACCAGAG (R)

Note: (F) = Forward primer, (R) = Reverse primer

Table 2.3. List of Oligonucleotides used for PCR Cloning in this Study

Construct	Oligonucleotide Sequences (5' → 3')	Template
E1A End	GCGGAATTCCCTCACCTTTCCCGGCAG (F) CGTGTCGACTTAACCCTCTTCATCCTCGTCGTCAC (R)	N/A
E1A M1	GGGTCCGGTTGCTATGCCAAAC (F) GTTTGGCATAGCAACCGGACCC (R)	pEGFP 13S E1A
E1A M2	GTCCGGTTTCTGCGCCAAACCTTG (F) CAAGGTTTGGCGCAGAAACCGGAC (R)	pEGFP 13S E1A
E1A M3	TCCGGTTTCTATGGCAAACCTTGTACC (F) GGTACAAGGTTTGCCATAGAAACCGGA (R)	pEGFP 13S E1A
E1A M4	TTTCTATGCCAGCCCTTGTACCGGAGG (F) CCTCCGGTACAAGGGCTGGCATAGAAA (R)	pEGFP 13S E1A
E1A M5	GGTTTCTATGCCAAACGCTGTACC (F) GGTACAGCGTTTGGCATAGAAACC (R)	pEGFP 13S E1A
E1A M6	CCAAACCTTGCTCCGGAGGTG (F) CACCTCCGGAGCAAGGTTTGG (R)	pEGFP 13S E1A
E1A M7	CAAACCTTGTAGCGGAGGTGATCGATC (F) GATCGATCACCTCCGCTACAAGGTTTG (R)	pEGFP 13S E1A
E1A M8	CCTTGTACCGGCGGTGATCGATC (F) GATCGATCACCGCCGGTACAAGG (R)	pEGFP 13S E1A
E1A M9	CTTGTACCGGAGGCGATCGATCTTACC (F) GGTAAGATCGATCGCCTCCGGTACAAG (R)	pEGFP 13S E1A
E1A M10	CCGGAGGTGGCAGATCTTACCTGC (F) GCAGGTAAGATCTGCCACCTCCGG (R)	pEGFP 13S E1A
E1A M11	TATGGCAAACGCTGTAGCGGAG (F) CTCCGCTACAGCGTTTGCCATA (R)	pEGFP 13S E1A
E1A M12	TCCGGTTTCTATCCCAAACGCTGTAC (F) GTACAGCGTTTGGGATAGAAACCGGA (R)	pEGFP 13S E1A
E1A M13	TCCGGTTTCTATGGCAAACGCTGTAC (F) GTACAGCGTTTGCCATAGAAACCGGA (R)	pEGFP 13S E1A
E1A M14	GTCCGGTTTCTATCGCAAACGCTGTAC (F) GTACAGCGTTTGCGATAGAAACCGGAC (R)	pEGFP 13S E1A
E1A M15	AACGCTGTAGCGGAGGTGATCG (F) CGATCACCTCCGCTACAGCGTT (R)	pEGFP 13S E1A

Table 2.3. List of Oligonucleotides used for PCR Cloning in this Study (Continued)

pEGFP CR2-CR3 E1A	GCGGAATTCCCTCACCTTTCCCGGCAG (F) ATTGTCGACTTAGGTAGGTCTTGCAGGCTC (R)	pEGFP 13S E1A
pcDNA3 HA BS69 Full Length	AATGAATTCATGCACCCTAAAGAGACCACCCGT (F) ATTCTCGAGTCATCTTTTCCGCGCGCAGGT (R)	pRc/CMV BS69
pcDNA3 HA BS69 No MYND	AATGAATTCATGCACCCTAAAGAGACCACCCGT (F) AATCTCGAGTCACGTGGGTATTTCCTGGCTAGAACTT ACTGCT (R)	pRc/CMV BS69

Note: (F) = Forward primer, (R) = Reverse primer

of purified PCR products were then combined and underwent a second round of PCR to generate the full length insert using Phusion Polymerase (Thermo Fisher) according to the manufacturer's protocol. 30ng of each PCR product was used in a 45µl reaction. This reaction underwent 15 cycles with the annealing and extension steps completed at 60^{o}C for 45 seconds and 72^{o}C for 60 seconds respectively. End primers were then added to the PCR mixture to a final concentration of 10µM each, and the reaction underwent 15 more cycles with the annealing and extension steps completed at 55^{o}C and 72^{o}C respectively for 60 seconds at each step. The PCR mixture was purified again using the GeneJET PCR Purification Kit (Thermo Fisher) according to the manufacturer's protocol. The purified insert and 2µg of the pBAIT backbone vector were then digested using EcoRI-HF and SalI-HF (New England Biolabs) at 37^{o}C for 1 hour in CutSmart buffer (New England Biolabs). After digestion, the purified insert and backbone vector were resolved on an agarose gel. The bands of interest were then cut and processed using EZ-10 Spin Columns (BioBasic) according to the manufacturer's protocol. The inserts were ligated into the corresponding sites of pBAIT using the Quick Ligation Kit (New England Biolabs) according to the manufacturer's protocol.

Plasmids (16-30) were generated using overlap extension (Ho et al., 1989). Polymerase chain reaction (PCR) amplification of E1A mutants was completed using Phusion Polymerase (Thermo Fisher) according to the manufacturer's protocol. The first round of PCR was completed using 10µM each of the matching forward and reverse primers (Table 2.3), and 250ng template DNA in a 50µl reaction. The reaction underwent 35 cycles with the annealing and extension steps completed at 60^{o}C for 30 seconds and 72^{o}C for 15 seconds respectively. The PCR mix was then incubated with 1µl DpnI (New England Biolabs) at 37^{o}C for 1 hour to digest the plasmid template before the PCR products were purified using the GeneJET PCR Purification Kit (Thermo Fisher) according to the manufacturer's protocol. Matching pairs of purified PCR products were then combined and underwent a second round of PCR to generate the full length insert using Phusion Polymerase (Thermo Fisher) according to the manufacturer's protocol. 30ng of each PCR product was used in a 45µl reaction. This reaction underwent 15 cycles with the annealing and extension steps completed at 60^{o}C for 45 seconds and 72^{o}C for 60 seconds respectively. End primers were then added to the PCR mixture to a final concentration of 10µM each,

and the reaction underwent 15 more cycles with the annealing and extension steps completed at 55^{o}C and 72^{o}C respectively for 60 seconds at each step. The PCR mixture was purified again using the GeneJET PCR Purification Kit (Thermo Fisher) according to the manufacturer's protocol. The purified insert and 2µg of the pBAIT backbone vector were then digested using EcoRI-HF and SalI-HF (New England Biolabs) at 37^{o}C for 1 hour in CutSmart buffer (New England Biolabs). After digestion, the purified insert and backbone vector were resolved on an agarose gel. The bands of interest were then cut and processed using EZ-10 Spin Columns (BioBasic) according to the manufacturer's protocol. The inserts were ligated into the corresponding sites of pBAIT using the Quick Ligation Kit (New England Biolabs) according to the manufacturer's protocol.

Plasmids (37, 39, 41) were generated using PCR to amplify the gene of interest from a plasmid template (Table 2.1). PCR amplification of the BS69 and E1A inserts using matching forward and reverse primers (Table 2.3) were completed using Phusion Polymerase (Thermo Fisher) according to the manufacturer's protocol. The BS69 template was provided by Dr. Rene Bernards (Utrecht University). PCR was completed using 10µM of forward and reverse primers each, and 250ng template DNA in a 50µl reaction. The reaction underwent 35 cycles with the annealing and extension steps completed at 60^{o}C for 30 seconds and 72^{o}C for 15 seconds respectively. The PCR mix was purified using the GeneJET PCR Purification Kit (Thermo Fisher) according to the manufacturer's protocol. The purified insert and the pcDNA3 HA or pEGFP backbone vector were then digested using a combination of EcoRI-HF and SalI-HF or XhoI (New England Biolabs) (Table 2.3) at 37^{o}C for 1 hour in CutSmart buffer (New England Biolabs). After digestion, the DNA was resolved on an agarose gel and the bands of interest were then cut and processed using EZ-10 Spin Columns (BioBasic) according to the manufacturer's protocol. The inserts were ligated into the corresponding sites of pcDNA3 HA or pEGFP using the Quick Ligation Kit (New England Biolabs) according to the manufacturer's protocol.

Plasmids 1-7, 20, 31-34 35-36, 38, 44-46, and 49-51 (Table 2.1) were previously generated by other members of the lab. Plasmids 1, 38 were generated by Dr. Joe Mymryk (University of Western Ontario); 2-7, 35 by Jennifer Curran; 20, 46 by Dr. Greg Fonseca (University of California San Diego); 31, 49 by Dr. Nik Avvakumov (Laval University); 32, 36 by Dr.

Ahmed Yousef (Masdar Institute of Science and Technology); 33 by Dr. Michael Shuen (University of Western Ontario); 34 by Stephanie Derbyshire; 44 by Cason King; 45 by Dr. Fred Dick (University of Western Ontario); and 50-51 by Dr. Jailal Ablack (University of California San Diego). Plasmids 40, 42, 43, and 52 were generated by subcloning previously generated inserts into a new backbone using the same restriction sites. Plasmid 47 was obtained from Clontech, 48 from Dr. Joe Torchia (University of Western Ontario) and 53 from Dr. Rene Bernards (Utrecht University).

All plasmids were transformed into competent DH5α *Escherichia coli* grown in lysogeny broth (LB) (10g/L tryptone, 5g/L yeast extract, and 86mM NaCl; all from Bioshop) supplemented with ampicillin (Amp) (50μg/ml, BioShop) or kanamycin (20μg/ml, BioShop) as required (Chung et al., 1989). Small-scale preparations of plasmid DNA used for cloning and yeast transformations were generated using the PureLink Quick Plasmid Miniprep Kit (Invitrogen), while large-scale preparations of plasmid DNA used for transfections were generated using the PureLink HiPure Plasmid Filter Midiprep Kit (Invitrogen) according to the manufacturer's protocol.

2.2 Yeast Culture and Transformation

Yeast culture and transformation were completed as previously described (Adams et al., 1997; Gietz et al., 1995). W303-1A (*MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) yeast were streaked from frozen stock onto yeast extract peptone (YEP) plates (yeast extract 10g/L, peptone 20g/L, 2% glucose, and 2% agar; all from Bioshop). The plates were incubated in 30^{o}C for 72 hours before colonies were picked for overnight growth with shaking at 30^{o}C in liquid YEP culture media (yeast extract 10g/L, peptone 20g/L, and 2% glucose; all from Bioshop). The liquid yeast culture was then diluted 1:2 in liquid YEP culture media and grown until the optical density at 600nm (OD_{600}) reached 1.0. The yeast cultures were then aliquoted into 1.5ml fractions and pelleted in a centrifuge for 15 seconds at room temperature. The pellets were resuspended in 1ml of 100mM lithium acetate (LiAc) (BioShop) and incubated for 5 minutes at 30^{o}C. After the incubation, the yeast mixture was pelleted in a centrifuge and resuspended in a 351μl transformation

mixture (34% polyethylene glycol 3350 (PEG), 100mM LiAc, and 0.14mg/ml salmon sperm DNA (SS-DNA); SS-DNA is from Sigma, all other components are from BioShop) with 200ng of each bait (plasmids 2-30, Table 2.1), prey (plasmids 31-32, Table 2.1) and reporter (plasmid 33, Table 2.1) plasmids. The suspension was then incubated for 30 minutes at 42^{o}C. Finally, the mixture was pelleted in a centrifuge, resuspended in 100µl ddH_2O, and plated onto synthetic drop-out plates (1.4g/L Y2001 Yeast Synthetic Drop-out Medium Supplements, 58µM L-tryptophan, 70mM D-glucose, 2% agar; Y2001 supplements are from Sigma, all other components are from Bioshop). The plates were then incubated in 30^{o}C for 72 hours before being picked for inoculation into 5ml of synthetic drop-out liquid media (1.4g/L Y2001 Yeast Synthetic Drop-out Medium Supplements, 58µM L-tryptophan, 70mM D-galactose; Y2001 supplements are from Sigma, all other components are from Bioshop) for overnight growth at 30^{o}C in a rotating drum. Galactose was used as the carbon source in place of glucose to induce expression of the GAL1 driven pJG4-5+ vector. The cultures were then used for yeast two-hybrid assays (Section 2.3) or protein extraction for western blot analysis (Section 2.4). Centrifugation was carried out at 15000 relative centrifugal force (RCF) unless otherwise stated.

2.3 Yeast Two-Hybrid Assays

Yeast two-hybrid assays were completed as previously described (Adams et al, 1997). 1.5ml of the overnight yeast cultures were pelleted in a centrifuge for 15 seconds at room temperature, washed, and resuspended in 1ml Z-Buffer (60mM Na_2HPO_4, 40mM Na_2H_2PO4, 10mM KCl, 1mM $MgSO_4$, and 50mM β-mercaptoethanol (BME); all from Bioshop, pH 7.0). The suspensions were then concentrated or diluted in Z-buffer until the OD_{600} was between 0.2 and 0.8. Then, 500µl of the suspensions were transferred to another microfuge tube containing 500µl Z-Buffer. Next, 20µl of 0.1% sodium dodecyl sulfate (SDS) (BioShop) and chloroform (Fisher Scientific) were added to each tube before being vortexed vigorously for 30 seconds. The samples were incubated in 30^{o}C for 15 minutes, then 200µl ortho-nitrophenyl-β-galactoside (ONPG) (4mg/ml, from Bioshop, in Z-Buffer) was added to each tube. The samples were then vortexed vigorously for 30 seconds before being incubated in 30^{o}C. To terminate the reaction, 0.5ml of 1M Na_2CO_3 (Bioshop) was

added to the samples after they turned pale yellow. The duration between the addition of ONPG and termination of the reaction was noted to calculate β-galactosidase activity. The samples were then centrifuged for 10 minutes to remove cellular debris and the optical density at 420nm (OD_{420}) was measured.

The equation to determine the β-galactosidase activity is as follows:

$$\text{Activity in Miller Units} = \frac{OD_{420}}{OD_{600} \times volume \times time} \times 1000$$

Where *volume* is the amount in ml of yeast suspension transferred to the new tube containing Z-buffer after the OD_{600} was measured; OD_{420} was measured after the addition of Na_2CO_3 and *time* is the period in minutes between the addition of ONPG and Na_2CO_3.

2.4 Yeast Protein Extraction of Western Blot Analysis

Yeast protein extractions were prepared as previously described (von der Haar, 2007). First, 1ml of liquid overnight yeast cultures were pelleted in a centrifuge for 15 seconds at room temperature, washed in ddH_2O, and resuspended in 200µl yeast lysis buffer (0.1M NaOH, 0.05M EDTA, 2% SDS, and 2% BME; all from BioShop). The suspensions were then incubated in 90^{o}C for 10 minutes. Next, 5µl of 4M acetic acid (BioShop) was added to the suspensions before the samples were briefly vortexed. The samples were then incubated again in 90^{o}C for 10 minutes. Finally, 50µl of yeast loading buffer (0.25M Tris-HCl pH 6.8, 50% glycerol, and 0.05% bromophenolblue, Tris-HCl and glycerol are from BioShop, bromophenolblue is from Sigma) was added to the suspensions before being vortexed and loaded into polyacrylamide gels for western blot analysis.

2.5 Cell Culture and Transfection

HT1080 human fibrosarcoma cells were used for the co-IP and luciferase experiments, while A549 lung epithelial carcinoma cells and IMR-90 primary lung fibroblast cells were used for the silencing RNA (siRNA) knockdown experiments. HT1080 cells were chosen because they transfect with high efficiency, while A549 and IMR-90 cells were used for their susceptibility to adenovirus infection for future experiments. All cells were cultured in Dulbecco's modified eagle medium (DMEM) (Multicell) supplemented with 10% fetal bovine serum (FBS) (Multicell), 100IU/ml penicillin (Multicell), and 100µg/ml streptomycin (Multicell). All cells were grown in a 37^{o}C incubator with 5% CO_2.

For co-IP experiments, HT1080 cells were grown on 10cm cell culture plates (Sarstedt). 2.2x10^{6} HT1080 cells were seeded per 10cm plate and transfected 24 hours later with 4µg each of plasmids containing HA-tagged BS69 (plasmid 39, Table 2.1), and a green fluorescent protein (GFP)-tagged E1A construct (plasmids 35-37, Table 2.1) or GFP alone (plasmid 34, Table 2.1). Transfections were carried out using X-tremeGene HP DNA Transfection Reagent (Roche) according to the manufacturer's protocol.

For luciferase assays and transfections to verify antibody specificity, HT1080 cells were grown on 6-well plates (Sarstedt). 3x10^{5} cells were seeded per well in the 6-well plates and transfected 24 hours later with 0.05µg of plasmids containing E1A constructs fused to GAL4 DNA binding domain (DBD) (plasmids 49-52, Table 2.1), or an empty GAL4 DBD vector (plasmid 47, Table 2.1); increasing concentrations of HA-tagged BS69 constructs (plasmids 39-41, Table 2.1); 0.5µg of luciferase reporter vector (plasmid 48, Table 2.1); and an empty pcDNA3 HA vector as required (plasmid 38, Table 2.1) to ensure a total amount of 2µg DNA transfected per well. Transfections were carried out using X-tremeGene HP DNA Transfection Reagent (Roche) according to the manufacturer's protocol.

2.6 Co-Immunoprecipitation

HT1080 cells were harvested 24 hours post-transfection using a cell scraper. The cells were pelleted in a centrifuge at 500 RCF for 2 minutes at 4^{o}C. The cells were washed with 5mL phosphate buffered saline (PBS) (173mM NaCl, 2.7mM KCl, 4.2mM Na_2HPO_4, and 1.5mM KH_2PO_4; all from BioShop), pelleted in a centrifuge at 500RCF, then resuspended in 1ml NP-40 lysis buffer (0.5% NP-40, 150mM NaCl, and 50mM Tris-HCl; all from BioShop, pH 7) supplemented with 0.5% Protease Inhibitor Cocktail (Sigma). The cells underwent lysis on ice for 10 minutes before being processed for 10 minutes in a centrifuge at 4^{o}C to pellet cell debris. After centrifugation, the lysates were transferred to new centrifuge tubes with 20µl of lysate aliquoted to another tube to be used as a 2% input control. 1µl of anti-GFP antibody (Table 2.4) and 120µl of a 10% Sepharose-Protein A slurry (Sigma) were added to the remaining 980µl of cell lysate. The samples were then incubated for 2 hours at 4^{o}C with gentle rocking. After the incubation period, the samples were pelleted at 500RCF in a centrifuge for 30 seconds at room temperature and washed with NP-40 lysis buffer. The washing process was repeated ten times to reduce nonspecific background binding. The Sepharose-antibody-protein complexes were resuspended in 25µl of 2X loading dye (49% NuPAGE LDS Sample Buffer from Thermo Fisher supplemented with 0.21M DTT from BioShop), while 10µl of 3X loading dye (70% NuPAGE LDS Sample Buffer from Thermo Fisher supplemented with 0.3M DTT from BioShop) was added to the 2% input controls. The input controls and Sepharose-antibody-protein complexes were then centrifuged for 30 seconds at room temperature before being incubated for 5 minutes at 98^{o}C to denature and separate proteins from the Sepharose beads. Samples were then loaded into polyacrylamide gels for western blot analysis.

2.7 Western Blot Analysis

Samples were loaded into NuPAGE 4-12% Bis-Tris polyacrylamide gradient gels (Invitrogen), resolved by electrophoresis at 200V, then transferred to a polyvinylidene fluoride (PVDF) membrane (Amersham Hybond from GE Healthcare Life Sciences) using

Table 2.4. List of Antibodies Used in this Study

Target	Animal of Origin	Usage	Dilution Factor from Stock	Company	Catalogue Number
LexA	Rabbit polyclonal	WB Primary	1:10000	Millipore	06-719
HA	Rat monoclonal	WB Primary	1:2000	Roche	11867423001
G6PD	Rabbit polyclonal	WB Primary	1:100000	Sigma	A-9521
GST	Rabbit polyclonal	WB Primary	1:5000	Sigma	G7781
6xHis	Rabbit polyclonal	PA Primary	1:5000	Abcam	ab1187
GFP	Rabbit polyclonal	Co-IP	1:1000	Clontech	632592
GFP	Rabbit polyclonal	WB Primary	1:2000	Clontech	632592
BS69	Rabbit monoclonal	WB/PA Primary	1:4000	Abcam	EP18343
Actin	Rabbit polyclonal	WB Primary	1:2000	Sigma	A-2066
Tubulin	Mouse monoclonal	WB Primary	1:5000	Sigma	T6199
Rat IgG	Goat polyclonal	WB Secondary	1:200000	Pierce	31470
Rabbit IgG	Goat polyclonal	WB/PA Secondary	1:10000	Jackson Laboratories	111-035-003
Mouse IgG	Goat polyclonal	WB Secondary	1:10000	Jackson Laboratories	115-035-003

Note: Antibodies in bold are conjugated to HRP
WB = western blot, PA = peptide array, Co-IP = co-immunoprecipitation

the XCell SureLock system (Invitrogen) according to the manufacturer's protocol. We used MES (2-[N-morpholino]ethanesulfonic acid) SDS running buffer (50mM MES, 50mM Tris, 3.47mM SDS, and 1.03mM EDTA; all from BioShop) during the gel electrophoresis step, and an in-house transfer buffer (25mM Bicine, 25mM Bis-Tris, 1.03mM EDTA, 20nM Chlorobutanol, and 10% Methanol; Bicine is from BioBasic, all other components are from BioShop) during the transfer step. The membranes were briefly soaked in tris buffered saline with Tween-20 (TBS-T) (20mM Tris, 136mM NaCl, and 0.1% Tween-20; Tween-20 is from Sigma, all other components are from BioShop) before being blocked at room temperature with shaking for one hour in 20ml of blocking buffer (5% w/v Skim Milk Powder from BioShop in TBS-T) before being incubated at 4^{o}C with shaking overnight in 20ml of primary antibody (Table 2.4) diluted in blocking buffer. The membranes were washed three times for 10 minutes each in 20ml of TBS-T. The membranes were then incubated at room temperature for 30 minutes with shaking in 20ml of species specific, horseradish peroxidase (HRP)-conjugated secondary antibody (Table 2.4) diluted in blocking buffer. The membranes were washed three more times for 10 minutes each in 20ml of TBS-T before using Luminata Crescendo Western HRP Substrate (Millipore) to detect protein according to the manufacturer's protocol. Images were developed on Amersham Hyperfilm enhanced chemiluminescence (ECL) membrane (GE Healthcare Life Sciences) using an automated film processor (Konica Minolta SRX-101A) according to the manufacturer's protocol.

2.8 Bradford Assay

Protein standards were created using 1, 2, 5, 10, 20, and 40µg/ml bovine serum albumin (BSA) (Sigma) in 800µl of ddH_2O. Samples to be measured were diluted 1:100 in ddH_2O to a final volume of 800µl. Next, 200µl of Protein Assay Dye Reagent (Bio-Rad) was added to the protein standards and samples (Bradford, 1976). After thorough mixing, 200µl of the solutions were loaded per well in duplicate into 96 well plates (Thermo Fisher). The 96 well plates were then read using the Multiskan Ascent plate reader (Thermo Fisher) according to the manufacturer's protocol. Data analysis was performed using Ascent Software (version 2.6, Thermo Fisher).

2.9 Luciferase Assay

HT1080 cells were harvested 24 hours post-transfection using a cell scraper. The cells were pelleted in a centrifuge at 500RCF for 2 minutes at room temperature and washed with 1ml of PBS. The cells were then resuspended in 200µl of Luciferase Cell Culture Lysis Reagent (Promega). The cells underwent lysis on ice for 10 minutes before being processed in a centrifuge for 10 minutes at 4^{o}C to pellet cell debris. After centrifugation, the lysates were transferred to new centrifuge tubes. Next, 20µl of lysates were aliquoted to verify protein expression using western blot analysis (Section 2.7), 8µl of lysates were aliquoted to measure total protein concentration using a Bradford Assay (Section 2.8), and 50µl of lysates were transferred to polystyrene test tubes (Fisher) to detect luciferase production. To detect luciferase production, 50µl of Luciferase Assay Substrate (Promega) was added to each sample immediately prior to light detection by the Lumat LB 9507 luminometer (Berthold). Measuring time was set to 10 seconds. Luciferase activity was first normalized to total protein concentration, then calculated as fold increase over that of cells transfected with an empty pM vector in place of a pM vector containing an E1A construct (Table 2.1).

2.10 siRNA Knockdown

Cells were seeded at a concentration of 3x10^{5} cells per well in 6 well plates. IMR-90 cells were transfected 24 hours after being seeded, while A549 cells were transfected 4 or 24 hours after being seeded. Downregulation of BS69 was performed using siRNA from Thermo Fisher (Silencer Select siRNA ID: s21153 and s21154) or Dharmacon (ON-TARGETplus SMARTpool). Downregulation of protein kinase A (PKA) Type 1a regulatory subunit was performed using siRNA from Thermo Fisher (Silencer Select siRNA ID: s286). The final concentration of siRNA transfected was 10nM or 20nM. Transfections were performed using siLentFect Lipid Reagent (BioRad) according to the manufacturer's protocol, using either 2µl or 4µl of the lipid transfection reagent per well for the 10nM and 20nM transfections respectively. Scrambled siRNA (Silencer Negative Control No.2, Ambion) was used as a negative control. Bradford Assays (Section 2.8) were used to ensure equal loading for western blot analysis.

2.11 Quantitative Reverse Transcription PCR

A549 cells were harvested 48 hours post-transfection using a cell scraper. Total RNA was extracted using the PureLink RNA Mini Kit (Thermo Fisher) and the PureLink DNase Set (Thermo Fisher) according to the manufacturer's protocol. cDNA was obtained by reverse transcribing 1µg of purified RNA using the SuperScript VILO cDNA Synthesis Kit (Thermo Fisher) according to the manufacturer's protocol. cDNA quantification via quantitative PCR (qPCR) was completed using the Power SYBR Green PCR Master Mix kit (Thermo Fisher) according to the manufacturer's protocol with 75ng cDNA per sample and 10µM oligonucleotide primers (BS69 F: GCAACACAGCACAAGCAACT, BS69R: GTACATGGCCTCCTCCTCAC). Measurements were taken on the QuantStudio 5 Real-Time PCR System (Thermo Fisher). Results were analyzed using the $2^{-\Delta\Delta CT}$ method, using glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the internal control and normalizing data to that of cells transfected to scrambled siRNA (Livak & Schmittgen, 2001; Vandesompele et al., 2002). No template and no reverse transcription controls were used to assess the presence of primer dimers and DNA contamination.

2.12 Purification of 6xHis and GST Tagged Proteins

Protein purification was completed using BL21 (DE3) RIL *Escherichia coli*. Cells transformed with pGEX4T1 or pET28a vectors (Table 2.1) were grown overnight in 5ml of LB supplemented with either chloramphenicol (Cam) (68µg/ml) and Amp (50µg/ml), or Cam (68µg/ml) and kanamycin (Kan) (20µg/ml) respectively. The overnight cultures were then diluted into 500ml of LB supplemented with antibiotics and grown with shaking at 37^{o}C until OD600 reached 0.5. To induce exogenous protein production, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1mM before incubating the cultures overnight with shaking at 16^{o}C. The cells were then pelleted in a centrifuge at 5500RCF at 4^{o}C for 10 minutes before being frozen at -80^{o}C overnight. The cells were then lysed using 20ml of an in-house lysis buffer (1mM phenylmethylsulfonyl fluoride (PMSF), 1mg/ml lysozyme, 2% Triton X-100, 100µM ZnCl; all from Bioshop) to purify glutathione S-transferase (GST)-tagged proteins, or lysis buffer supplemented with 20mM

imidazole (Bioshop) to purify 6xHis-tagged proteins. The lysates were then incubated with shaking at 37^{o}C for 1 hour before being sonicated three times for 15 seconds each at 50% of the microtip limit using a Vibracell sonicator (Sonics and Materials). The lysates were then centrifuged at 5500RCF at 4^{o}C for 10 minutes to pellet cell debris. 3ml of Nickel-NTA (Qiagen) or 3ml of Glutathione 4 Fast Flow (GE Healthcare Life Sciences) slurries were loaded into 20ml chromatography columns (Bio-Rad). The columns were then equilibrated using two 20ml washes of PBS. The cell lysates were then run through the corresponding chromatography columns. The glutathione columns were washed three times with 20ml of PBS per wash, while the nickel columns were washed three times with 20ml of PBS supplemented with 20mM imidazole per wash. Ten 1ml fractions were collected by eluting the columns with either 10mM glutathione (BioShop) in PBS, or 250mM imidazole in PBS. Samples were collected before IPTG was added, after overnight growth in IPTG, after lysate flow-through, and after wash steps to monitor protein expression and column affinity.

2.13 Coomassie Blue Staining

Coomassie blue staining of polyacrylamide gels were completed as previously described (Neuhoff et al., 1988). After electrophoresis, polyacrylamide gels were incubated with shaking at room temperature in Coomassie staining solution (50% methanol, 0.05% Coomassie Brilliant Blue R-250, 10% acetic acid; all from BioShop) for one hour. The gels were then rinsed with water before being incubated with shaking at room temperature in detain solution (30% methanol, 10% acetic acid; all from BioShop) and paper towels to draw up excess Coomassie Blue dye. Paper towels and destain solution were changed as needed until the gel background was clear of excess blue colouration.

2.14 Peptide Array

Peptide arrays were provided by the laboratory of Dr. Shawn Li (University of Western Ontario) and were synthesized using previously described methods (Frank, 2002; Huang

et al., 2008). Design for the positional scanning library was completed using the Genscript online peptide array design tool. Arrays were synthesized using Fmoc solid-phase peptide synthesis on amino-functionalized cellular membranes (Whatman) automated by the MultiPep RSI Peptide Synthesizer (Intavis) according to the manufacturer's protocol. After the array was synthesized, the peptides on the membranes were deprotected for 1.5 hours using a trifluoroacetic acid (TFA) cocktail solution (95% TFA, 3% triisopropylsilane (TIPS) in ddH_2O). The membranes were washed with dichloromethane (DCM) three times for 2 minutes each, then with dimethylformamide (DMF) three times for 2 minutes each, and finally with absolute ethanol three times for 2 minutes each. The membranes were then left overnight to air dry and stored in -20^oC. The membranes were reactivated by immersion in 50% ethanol for 15 minutes, followed by a 15-minute incubation in ddH_2O. The membranes were then blocked in 3% BSA at 4^oC overnight with shaking before being probed with 1μM 6xHis-tagged BS69 in 3% BSA at 4^oC overnight with shaking. The membrane was then washed with PBS three times for 10 minutes each and then incubated in HRP-conjugated anti-6xHis tag antibody (Table 2.4) diluted in 3% BSA for one hour. The membranes were washed three more times for 10 minutes in PBS before using Luminata Crescendo Western HRP Substrate (Millipore) to detect protein according to the manufacturer's protocol. The membrane was imaged using a Gel Doc XR (Bio-Rad).

To strip the membrane, the array was washed with shaking with 20ml H_2O three times for 10 minutes each. The membrane was then incubated with 20ml of stripping mix A (8M urea, 1% SDS in PBS) in a sonication bath at 40^oC three times for 10 minutes each. The membrane was then incubated with shaking using 20ml of stripping mix B (10% acetic acid, 50% ethanol in H_2O) three times for 10 minutes each, followed by incubation with shaking using 20ml of absolute ethanol three times for 10 minutes each. Finally, the membranes were washed with shaking using PBS three times for 5 minutes each. After this process, the membrane was blocked in 3% BSA at 4^oC overnight with shaking before being probed as necessary.

2.15 Homology Modelling

Protein sequences from all 21 ZMYND gene family members were obtained from Uniprot and aligned using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011). The phylogenetic tree was then visualized using FigTree (version 1.4.2). The closest relative of BS69 with a pre-existing structural data, which is ZMYND5, was used as the template for the homology model. The ZYMND5 structural data file (PDB ID: 4A24) was obtained from RCSB online repository and visualized using UCSF Chimera (version 1.9) (Kateb et al., 2013). The MODELLER plugin was used to construct the homology model (Webb & Sali, 2016).

3 Results

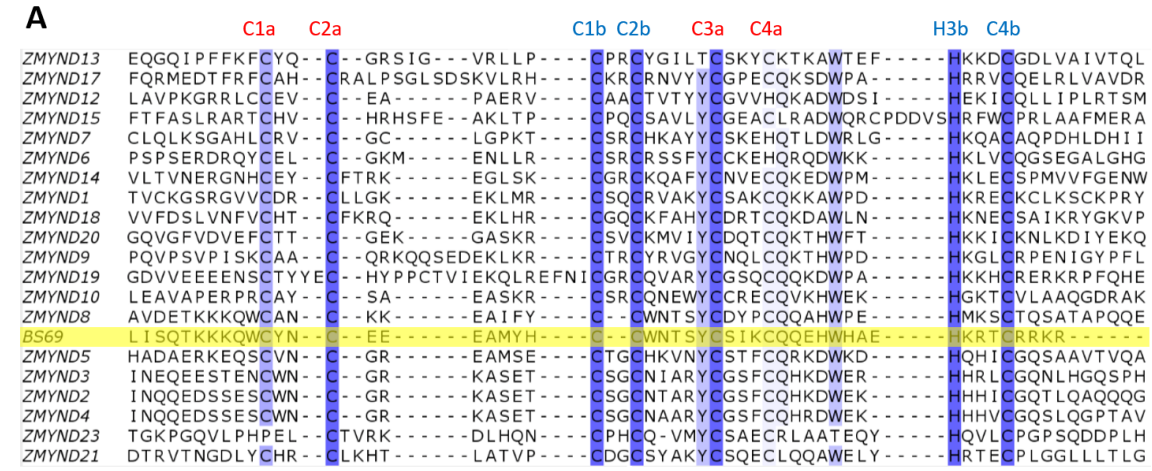
3.1 Mapping the Interaction Between E1A and BS69

3.1.1 Using Homology Modelling to Construct the MYND Domain of BS69

The cellular environment consists of complex and dynamic networks of protein-protein interactions. Each constituent in these networks plays a specific role in ensuring the functionality of a larger encompassing biochemical pathway. To fully understand the details of a specific protein-protein interaction, a model with atomic-level resolution is typically required. These models can be determined experimentally using several techniques including protein crystallography and nuclear magnetic resonance (NMR) spectroscopy. Unfortunately, both these techniques require highly purified protein and each have their own unique challenges and limitations as well: obtaining high quality crystals is difficult for crystallography, while NMR is limited to smaller proteins (Yee et al., 2005). An alternative to experimentally determining protein structures is to use *in silico* methods such as homology modelling and protein-ligand docking. A homology model of the BS69 MYND domain based on DEAF-1, or ZMYND5, was previously constructed by the Sattler group in 2013 (Kateb et al., 2013). Thus, we set out to re-create this homology model, and use it to analyze the interaction between E1A and the MYND domain of BS69 by using protein-ligand docking.

To construct a homology model of the MYND domain, we first performed a multiple sequence alignment of all the known ZMYND (ZMYND1-21, 23) family of proteins (Figure 3.1A). Protein sequences were obtained from UniProt and the multiple sequence alignment was completed using Clustal Omega (Goujon et al., 2010; Sievers et al., 2011). A phylogenetic tree was also created using Clustal Omega, which was then used to determine the closest relative of BS69 with a preexisting structure (Figure 3.1B). ZMYND5 was chosen as the template to create the homology model using the MODELLER plugin in UCSF Chimera (Webb & Sali, 2016). ZMYND5 is shown in the

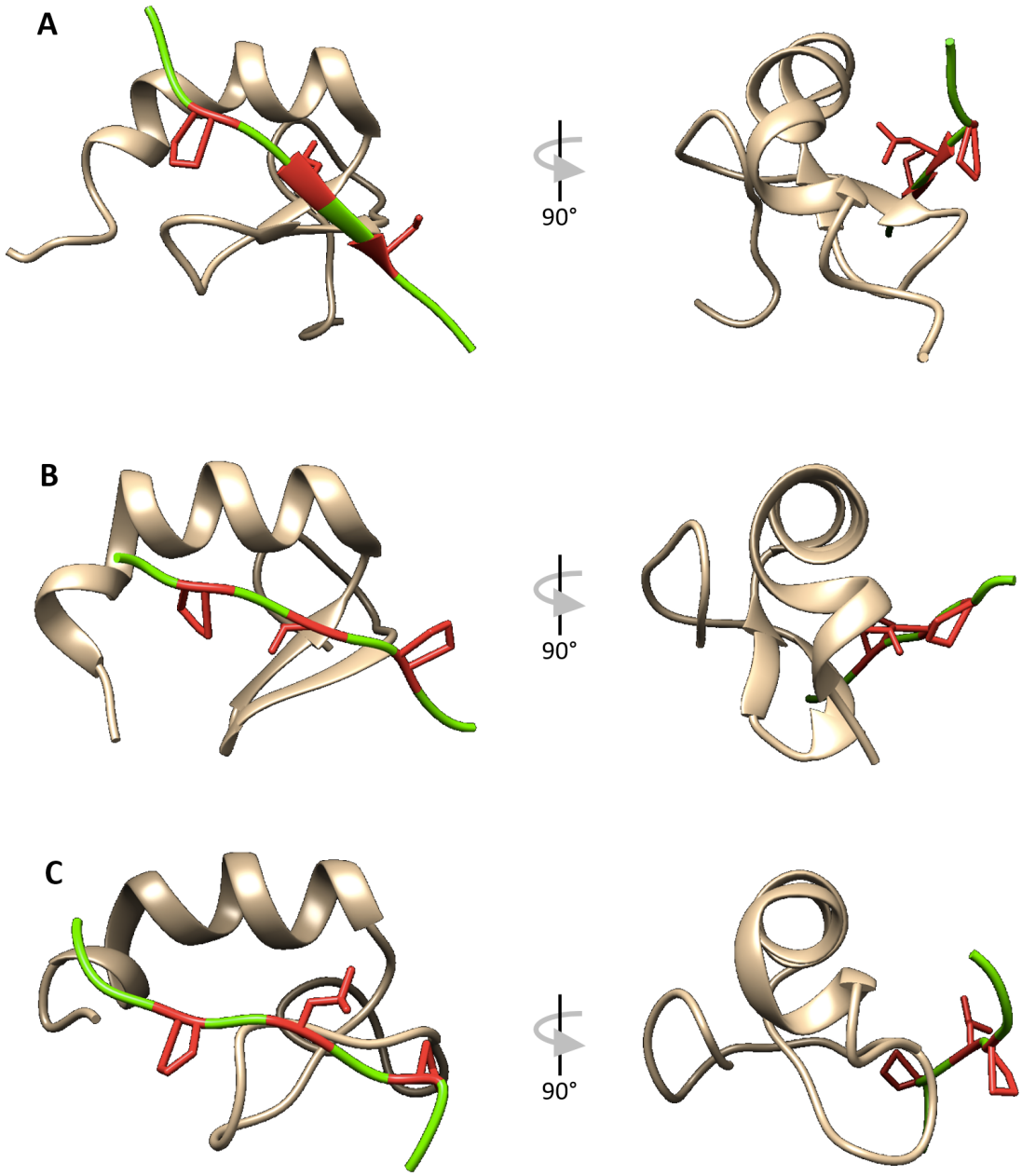
Figure 3.1. Homology model of the BS69 MYND domain. A) Multiple sequence alignment of the MYND domains of 21 proteins of the ZMYND family. BS69 (ZMYND11) is highlighted in yellow. The cysteine and histidine residues used to coordinate zinc ions are labelled in the same scheme as Figure 6. Residues with high conservation are shown in blue, with darker shading indicating higher levels of conservation. Sequence alignment was generated using Clustal Omega. B) Phylogenetic tree of the 21 members of the ZMYND family of proteins. BS69 is highlighted in blue. ZMYND5, immediately below BS69, was used as the template for the homology model. C) Ribbon structures of ZMYND5 (PDB: 4A24) and the BS69 homology model (Kateb et al., 2013). ZMYND5 is shown in silver in the first row, while the homology model is shown in yellow in the third row. An overlay of the two structures is shown in the second row. The structures on the left depict the side of the protein that interacts with E1A and NCOR2, while the structures on the right depict a 90° rotation along the Y axis. Zinc ions are shown as grey spheres. Side chains of the cysteine and histidine residues used to coordinate the zinc ions are shown. The sulfur and nitrogen atoms are coloured yellow and blue respectively. Ribbon structures of ZMYND5 adapted from (Kateb et al., 2013).


A
C1a
C2a
C1b
C2b
C3a
C4a
H3b
C4b
ZMYND13 EQGQIPFFKFCYQ--C--GRSIG---VRLLP----CPRCYGILTCSKYCKTKAWTEF-----HKKDCGDLVAIVTQL
ZMYND17 FQRMEDTFRFCAH--CRALPSGLSDSKVLRH----CKRCRNVYYCGPECQKSDWPA------HRRVCQELRLVAVDR
ZMYND12 LAVPKGRRLCCEV--C--EA------PAERV----CAACTVTYYCGVVHQKADWDSI-----HEKICQLLIPLRTSM
ZMYND15 FTFASLRARTCHV--C--HRHSFE--AKLTP----CPQCSAVLYCGEACLRADWQRCPDDVSHRFWCPRLAAFMERA
ZMYND7 CLQLKSGAHLCRV--C--GC------LGPKT----CSRCHKAYYCSKEHQTLDWRLG-----HKQACAQPDHLDHII
ZMYND6 PSPSERDRQYCEL--C--GKM-----ENLLR----CSRCRSSFYCCKEHQRQDWKK------HKLVCQGSEGALGHG
ZMYND14 VLTVNERGNHCEY--CFTRK------EGLSK----CGRCKQAFYCNVECQKEDWPM------HKLECSPMVVFGENW
ZMYND1 TVCKGSRGVVCDR--CLLGK------EKLMR----CSQCRVAKYCSAKCQKKAWPD------HKRECKCLKSCKPRY
ZMYND18 VVFDSLVNFVCHT--CFKRQ------EKLHR----CGQCKFAHYCDRTCQKDAWLN------HKNECSAIKRYGKVP
ZMYND20 GQVGFVDVEFCTT--C--GEK-----GASKR----CSVCKMVIYCDQTCQKTHWFT------HKKICKNLKDIYEKQ
ZMYND9 PQVPSVPISKCAA--C--QRKQQSEDEKLKR----CTRCYRVGYCNQLCQKTHWPD------HKGLCRPENIGYPFL
ZMYND19 GDVVEEEENSCTYYEC--HYPPCTVIEKQLREFNICGRCQVARYCGSQCQQKDWPA------HKKHCRERKRPFQHE
ZMYND10 LEAVAPERPRCAY--C--SA------EASKR----CSRCQNEWYCCRECQVKHWEK------HGKTCVLAAQGDRAK
ZMYND8 AVDETKKKQWCAN--C--KK------EAIFY----C--CWNTSYCDYPCQQAHWPE------HMKSCTQSATAPQQE
BS69 LISQTKKKQWCYN--C--EE------EAMYH----C--CWNTSYCSIKCQQEHWHAE-----HKRTCRRKR------
ZMYND5 HADAERKEQSCVN--C--GR------EAMSE----CTGCHKVNYCSTFCQRKDWKD------HQHICGQSAAVTVQA
ZMYND3 INEQEESTENCWN--C--GR------KASET----CSGCNIARYCGSFCQHKDWER------HHRLCGQNLHGQSPH
ZMYND2 INQQEDSSESCWN--C--GR------KASET----CSGCNTARYCGSFCQHKDWEK------HHHICGQTLQAQQQG
ZMYND4 INQQEDSSESCWN--C--GR------KASET----CSGCNAARYCGSFCQHRDWEK------HHHVCGQSLQGPTAV
ZMYND23 TGKPGQVLPHPEL--CTVRK------DLHQN----CPHCQ-VMYCSAECRLAATEQY-----HQVLCPGPSQDDPLH
ZMYND21 DTRVTNGDLYCHR--CLKHT------LATVP----CDGCSYAKYCSQECLQQAWELY-----HRTECPLGGLLLTLG
B
ZMYND13
ZMYND15
ZMYND23
ZMYND21
ZMYND7
ZMYND12
ZMYND17
ZMYND6
ZMYND9
ZMYND20
ZMYND14
ZMYND1
ZMYND18
ZMYND8
BS69
ZMYND5
ZMYND3
ZMYND2
ZMYND4
ZMYND10
ZMYND19
C
90°
90°
90°

top row in silver, the BS69 homology model is on the bottom row in yellow, and the middle row shows a superimposed image of both models (Figure 3.1C). The models are depicted using ribbon diagrams, with the zinc ions shown as grey spheres and the zinc-coordinating histidine and cysteine side chains protruding from the ribbon.

Protein-ligand docking was then completed using ClusPro (Kozakov et al., 2017). The PDB file of the BS69 MYND homology model was uploaded as the *receptor*, and a flexible peptide with the E1A derived sequence "SMPNLVPEVI" was created in UCSF Chimera and submitted as the *ligand*. The top row shows the interaction between a peptide from nuclear receptor co-repressor 2 (N-CoR2) and the DEAF-1 MYND domain (Figure 3.2A), the middle row shows the structure of an EBNA2 peptide interacting with the BS69 MYND domain (Figure 3.2B), and the bottom row shows the structure generated by protein-ligand docking of the BS69 MYND homology model with the submitted E1A peptide (Figure 3.2C) (Harter et al., 2016; Liu et al., 2007). The N-CoR2 peptide contains a PXLXS (sequence: TISNPPPLISSAK) motif located in the same spatial position as the PXLXP motif for EBNA2 (sequence: SMPELSPVL) and E1A (sequence: SMPNLVPEVI). The ligands are shown in green and the core motif is shown in red. The central leucine residue in the PXLXS and PXLXP core motifs of N-CoR2 and EBNA2 are directed inwards towards the MYND domain of DEAF-1 and BS69 MYND, but the leucine residue in E1A is directed away from BS69 MYND. However, the proline residues of the EBNA2 and E1A peptides share similar orientations. Thus, this homology model may be somewhat useful in helping visualize the E1A-BS69 protein-protein interaction and predicting residues that make contact between these two proteins. However, precise conclusions about the interaction mechanism should not be drawn by solely referring to this model and must be validated using other experimental approaches.

Figure 3.2. Protein-ligand docking of E1A and the MYND domain of BS69. NMR ribbon structures of A) NCOR2 peptide in complex with DEAF-1 MYND (PDB: 2ODD) and B) EBNA2 peptide in complex with BS69 MYND (PDB: 5HDA) are shown in the first two rows (Harter et al., 2016; Liu et al., 2007). C) Protein-ligand docking of an E1A peptide in complex with the homology model of BS69 MYND. The NCOR2 peptide consists of residues 1109-1121 with the sequence TISNPPPLISSAK, the EBNA2 peptide consists of residues 381-389 with the sequence SMPELSPVL, and the E1A peptide consists of residues 111-120 with the sequence SMPNLVPEVI. Peptides are shown in green, with the PXLXS or PXLXP motif shown in red. The MYND domains of the various proteins are show in beige. NCOR2 in complex with DEAF-1 MYND adapted from (Liu et al., 2007). EBNA2 peptide in complex with BS69 MYND adapted from (Harter et al., 2016).


A
90°
B
90°
C
90°

3.1.2 Determining Conservation of the Interaction between BS69 and E1A of Adenovirus Species A-F

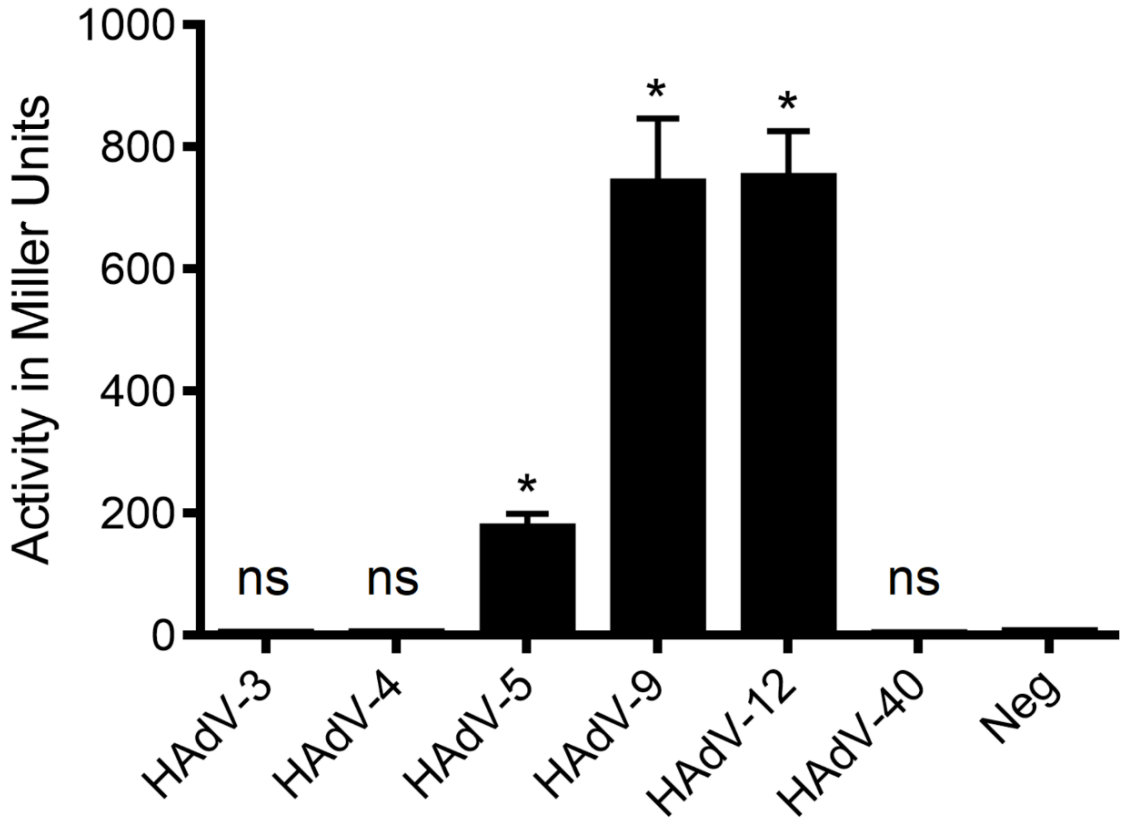
The interaction between E1A and BS69 was initially discovered through an E1A pulldown screen performed by the Bernards group in 1995 (Hateboer et al., 1995). Subsequent studies on this interaction have all been conducted using E1A from HAdV-5 (Ansieau & Leutz, 2002). Yet, interactions between E1A and other important host factors, such as pRb, p300, and PKA, are conserved amongst many other adenovirus species (Avvakumov et al., 2004; King et al., 2016; Pelka et al., 2008). Important functions of a protein are more likely to be conserved between different species, so we set out to determine if E1A from other adenovirus species can interact with BS69.

We completed these experiments using yeast two-hybrid assays. We used E1A from HAdV-3, 4, 5, 9, 12, and 40, which are representative types from adenovirus species B, E, C, D, A, and F respectively (Figure 3.3). Yeast expression vectors containing E1A CR2 from each adenovirus type in a bait vector, the MYND domain of BS69 in a prey vector, and a yeast two-hybrid reporter vector were transformed into W303-1A yeast. We chose to use constructs expressing only CR2 instead of full length E1A to reduce background activity, as the N-terminus and CR3 both facilitate intrinsic transcriptional activation by E1A (Ablack et al., 2010). The negative control consisted of yeast transformed with HAdV-5 E1A CR2, an empty prey plasmid, and the reporter vector. The yeast were incubated for 72 hours in selective agar plates before being picked for overnight growth in selective liquid media. The liquid cultures were then processed, and β-galactosidase activity in Miller units was calculated. Significance markers are assigned in comparison to the negative control (Figure 3.3).

The interaction between BS69 and E1A is conserved in HAdV-5, 9, and 12, but not HAdV-3, 4, and 40 (Figure 3.3A). The activity in Miller units for the interaction between BS69 and E1A from HAdV-5, 9, and 12 are significantly higher than the negative control by 28-, 118-, and 119-fold respectively. Interestingly, the PXLXP motif is present in E1A from all HAdV species that are able to interact with BS69, while absent in E1A from species

Figure 3.3. Conservation of the interaction between BS69 and E1A in different adenovirus species. A) Yeast two-hybrid assay of E1A CR2 of HAdV-3, 4, 5, 9, 12, and 40 in bait and BS69 MYND in prey in W303-1A yeast. β-galactosidase activity in Miller units is shown on the Y-axis, while the adenovirus species are indicated on the X-axis. The negative control (Neg) consists of yeast transformed with E1A CR2 of HAdV-5 in bait and an empty prey vector. Results are shown as mean ± SD, n=3. Significance markers are assigned in comparison to the negative control (ns = not significant; * $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test). B) Multiple sequence alignment of E1A CR2 of HAdV-3, 4, 5, 9, 12, and 40. Alignment was completed using Clustal Omega. The PXLXP motifs are outlined in red boxes. Residues with high conservation are shown in blue, with darker shading indicating higher levels of conservation.

A


Activity in Miller Units
1000
800
600
400
200
0
ns
ns
*
*
*
ns
HAdV-3
HAdV-4
HAdV-5
HAdV-9
HAdV-12
HAdV-40
Neg

B


HAdV-3 PPETLVTPGVVVESGIGGKKLPDLGAAE-MDLRCYEEGFPPSDDEDGETEQSIHTAVNEGVKAASDV
HAdV-4 LHTPRHDRGEKEIPGLKWEK-MDLRCYEECLPPSDDEDEQA---IQNAASHGVQAVSE
HAdV-5 -PHLSRQPEQPEQRALGPVSMPNLVPEV-IDLTCHEAGFPPSDDEDEEG
HAdV-9 ---LHTPTLSP---------IPELEEEDELDLRCYEEGFPPSDSEDERGEQTMALISDYACVIVEEQ
HAdV-12 -------VLSPVCEPIGGECMPQLHPED-MDLLCYEMGFPCSDSEDEQDENGMAHVSASAAAAAADREREE
HAdV-40 ------------SGSGDSGVGEELLPVDLDLKCYEDGLPPSDPETDEATEAEEEAAMPTY--VNENENELVLDCPENPG

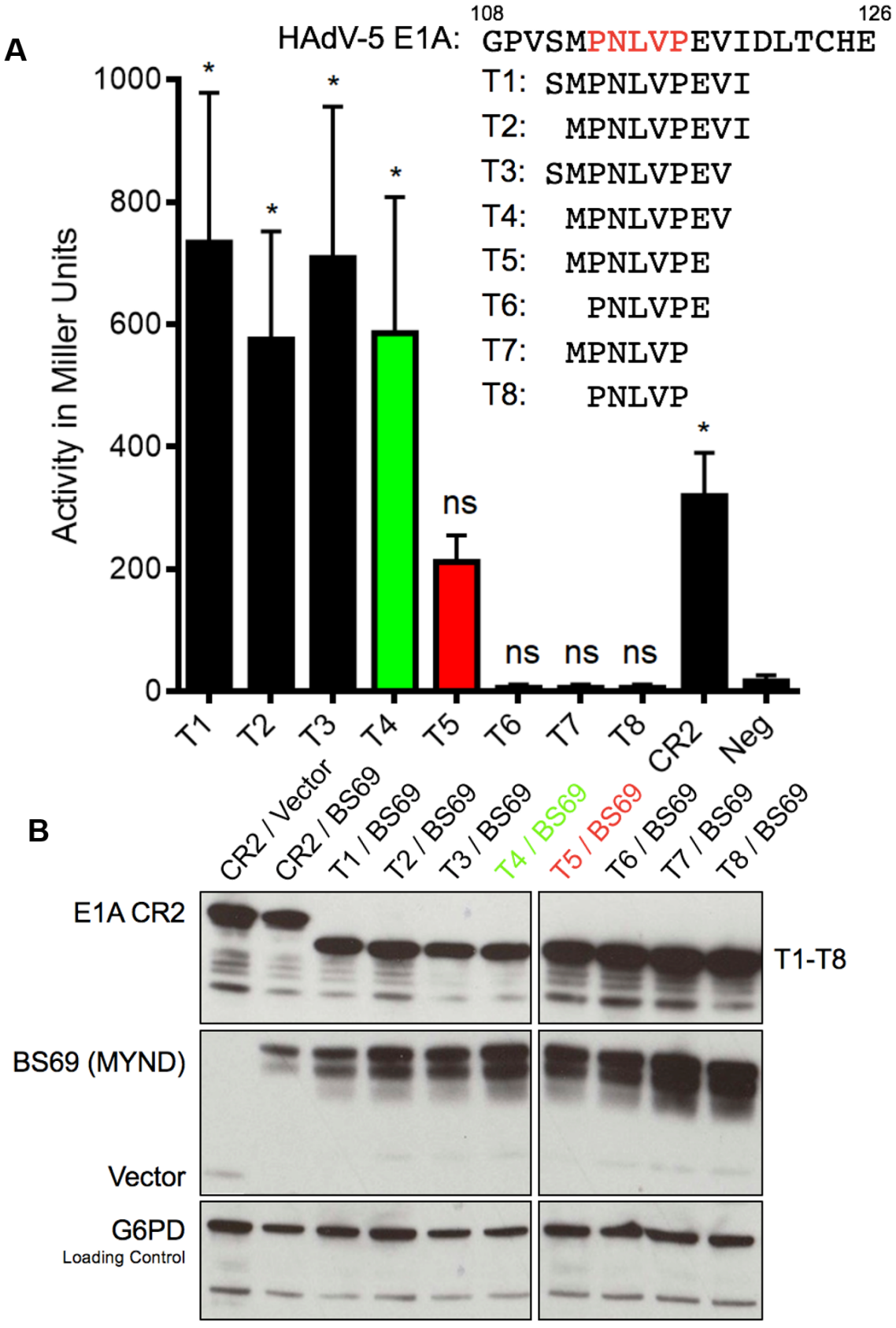
that are unable to interact with BS69 (Figure 3.3B). Because most molecular reagents used to study adenoviruses have been constructed using the HAdV-5 genome, the majority of this project will focus on E1A from HAdV-5.

3.1.3 Defining the Minimal Interaction Region on HAdV-5 E1A for BS69

The PXLXP motif was defined by the Leutz group in 2002 by sequence comparison of E1A and ENBA2, which is a protein from Epstein Barr Virus (EBV) that is also able to interact with BS69 (Ansieau & Leutz, 2002). This motif is present not only in viral proteins, but also on host proteins such as skeletal α-nascent polypeptide-associated complex (skNAC) and chromatin assembly factor 1 (CAF1), which use the PXLXP motif to mediate interactions with the MYND containing proteins m-Bop and C-C motif chemokine receptor 4 (CCR4) respectively (Chou et al., 2017; Sims et al., 2002). The core PXLXP motif on E1A is necessary but not sufficient to mediate a strong interaction with BS69. This phenomenon is also true for other E1A interactors, including the LXCXE-mediated interaction with pRb (DeCaprio, 2009). Determining the minimal interacting region allows us to better understand the mechanism of interaction, and would help improve the experimental design of our proposed peptide array experiment. This information will allow us to use the shortest possible peptide sequences for the array, leading to decreased cost and increased sequence fidelity of the synthesized peptides. Thus, we set out to determine the minimal number of flanking residues outside of the core PXLXP motif required for E1A to interact strongly with BS69.

This set of experiments was completed using yeast two-hybrid assays. We used a series of HAdV-5 E1A truncation mutants denoted T1 (SMPNLVPEVI, residues 111-120) through T8 (PNLVP, residues 113-117) in bait, the MYND domain of BS69 in prey, and a LacZ reporter vector transformed into W303-1A yeast (Figure 3.4). The negative control consisted of yeast transformed with E1A CR2, an empty prey plasmid, and the reporter vector. Transformed yeast were grown for 72 hours on selective agar plates before being picked for overnight growth in selective liquid media. The liquid cultures were then

Figure 3.4. Determining the minimal interacting region on E1A required to bind with BS69. A) Yeast two-hybrid assay of E1A truncation mutants in bait and BS69 MYND in prey in W303-1A yeast. Sequences of the E1A truncation mutants with the corresponding designations are shown above the graph. β-galactosidase activity in Miller units is shown on the Y-axis, while the truncation mutants are indicated on the X-axis. The negative control (Neg) consisted of yeast transformed with E1A CR2 in bait and an empty prey vector, while the positive control (CR2) consisted of yeast transformed with E1A CR2 in bait and BS69 MYND in prey. Results are shown as mean ± SD, n=4. Significance markers are assigned in comparison to the negative control (ns = not significant; * $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test). B) Western blot of yeast cell lysate to verify protein expression. Negative control is labelled as CR2 / Vector. Bait and prey proteins were visualized using anti-LexA DBD and anti-HA antibodies respectively. G6PD was used as a loading control.


A
HAdV-5 E1A: 108 GPVSMPNLVPEVIDLTCHE 126
T1: SMPNLVPEVI
T2: MPNLVPEVI
T3: SMPNLVPEV
T4: MPNLVPEV
T5: MPNLVPE
T6: PNLVPE
T7: MPNLVP
T8: PNLVP
Activity in Miller Units
1000
800
600
400
200
0
*
ns
T1
T2
T3
T4
T5
T6
T7
T8
CR2
Neg
B
CR2 / Vector
CR2 / BS69
T1 / BS69
T2 / BS69
T3 / BS69
T4 / BS69
T5 / BS69
T6 / BS69
T7 / BS69
T8 / BS69
E1A CR2
T1-T8
BS69 (MYND)
Vector
G6PD
Loading Control

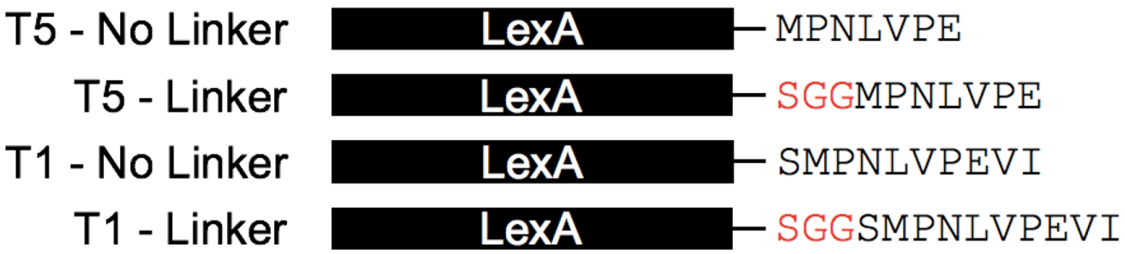
processed, and β-galactosidase activities in Miller units were calculated. Significance markers were assigned in comparison to the negative control (Figure 3.4A). Compared to the entire E1A CR2 fragment, the truncation mutant T4, indicated in green with the sequence MPNLVPEV, was the shortest mutant that interacted with BS69 at levels comparable to CR2 (Figure 3.4A). Subsequent removal of the C-terminal valine (T5) and the N-terminal methionine (T6) resulted in stepwise decreases in binding activity. T4 and T5, in red, showed a 35- and 12-fold increase in activity over the negative control respectively, while the T6 and further truncations were totally defective for binding in this assay. Notably, the CR2 positive control exhibited less activity than truncation mutants T1-T4 possibly due to steric hindrance or competitive binding with other E1A interactors in this region, such as UBC9. Western blots of yeast lysate ensured these observations were not simply due to reduced protein expression by the non-binding truncation mutant fusion proteins (Figure 3.4B). The division between the blots show the samples being run and probed concurrently on separate gels. Glucose-6-phosphate dehydrogenase (G6PD) was used as the loading control for both gels. I also performed a second set of experiments using constructs that included a flexible linker, with the sequence SGG, between the LexA DNA binding domain and the E1A truncation peptide (Figure 3.5A). The inclusion of this linker in the fusion protein resulted in increased β-galactosidase activity of approximately 100 Miller units for truncation mutants T1 and T5, but did not increase the β-galactosidase activity for a negative control where the yeast were transformed with an empty prey vector (Figure 3.5B).

E1A CR2 contains many overlapping interaction motifs used to bind to pRb, BS69, UBC9, the S2 component of the 19S regulatory complex of the proteasome, and more recently, STING (Lau et al., 2015; Pelka et al., 2008). Because of this, designing a specific mutant that abrogates binding with BS69 without affecting the neighbouring interacting partners is challenging. After finding the minimal interacting region, we set out to determine which residues within this region are important in binding to BS69. To accomplish this, we used another series of yeast two-hybrid assays.

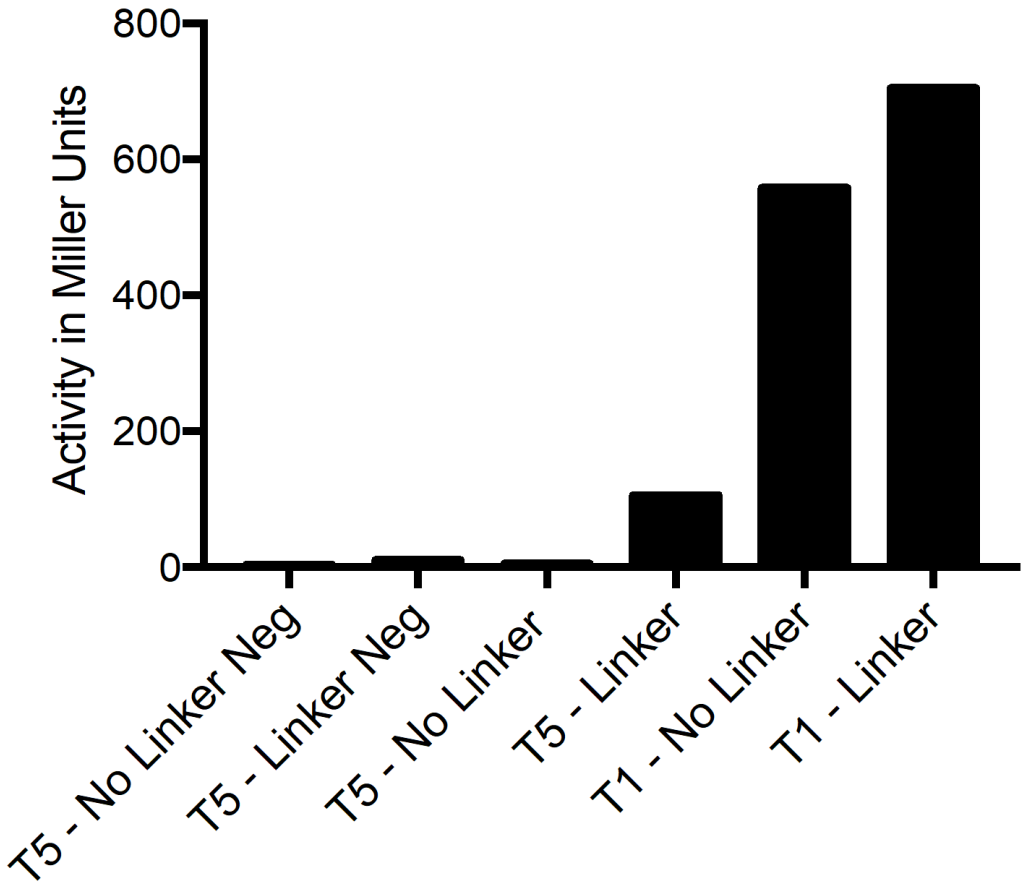
Alanine scanning mutants denoted M1 (E1A S111A) through M10 (E1A I120A) in the context of CR2 in bait, the MYND domain of BS69 in prey, and a reporter vector were

Figure 3.5. The effects of including a flexible SGG linker on the measurement of E1A-BS69 interactions using the yeast two-hybrid assay. A) Diagrammatic representation of E1A truncation mutants in bait with or without flexible linkers. Sequences of the T5 and T1 truncation mutants with or without SGG linkers are shown. B) Yeast two-hybrid assay of E1A truncation mutants, with or without SGG linkers, in bait and BS69 MYND in prey in W303-1A yeast. β-galactosidase activity in Miller units is shown on the Y-axis, while the truncation mutants are indicated on the X-axis. Negative controls (T5 – No Linker Neg and T5 – Linker Neg) consisted of yeast transformed with the respective bait proteins with an empty prey vector. Results are shown as values from a single experiment.

A


T5 - No Linker
LexA
MPNLVPE
T5 - Linker
LexA
SGGMPNLVPE
T1 - No Linker
LexA
SMPNLVPEVI
T1 - Linker
LexA
SGGSMPNLVPEVI

B


Activity in Miller Units
800
600
400
200
0
T5 - No Linker Neg
T5 - Linker Neg
T5 - No Linker
T5 - Linker
T1 - No Linker
T1 - Linker

transformed into W303-1A yeast (Figure 3.6A). The negative control consisted of yeast transformed with E1A CR2, an empty prey plasmid, and the reporter vector. The wildtype positive control was similar to the negative control, except with the MYND domain of BS69 in prey. The transformed yeast were grown for 72 hours on selective agar plates before being picked for overnight growth in selective liquid media. The liquid cultures were processed, and β-galactosidase activity in Miller units was calculated. Significance markers were assigned in comparison to the wildtype positive control (Figure 3.6B).

E1A constructs M3 and M5, which correspond to the P113A and L115A point mutations respectively, showed significantly decreased binding affinity to BS69, at 4- and 18-fold less respectively, compared to the wildtype control (Figure 3.6B). Both these mutants target the PXLXP core motif. Interestingly, the M7 construct, which corresponds to the P117A point mutant, did not show significantly different binding affinity to BS69 compared to the wildtype control, even though this mutant targets the C-terminal proline residue in the PXLXP motif. Yeast transformed with wildtype E1A CR2 in bait, with an empty prey vector was used as a negative control.

Concurrently with the alanine scanning mutation panel, we also constructed several double and triple point mutants to create an E1A construct that thoroughly abrogates binding with BS69. To help with the design of these point mutants, we referred to an NMR structure of an EBNA2 peptide in complex with the MYND domain of BS69 (Harter et al., 2016). We targeted M112, P113, L115, and P117 for mutation, as the side chains of these residues protrude towards and make direct contact with the BS69 MYND domain (Harter et al., 2016). Once again, we used a series of yeast two-hybrid assays by transforming E1A double and triple point mutants denoted M11 through M15 in the context of CR2 in bait, the MYND domain of BS69 in prey, and a reporter vector into W303-1A yeast (Figure 3.7A). All double and triple mutants showed significantly decreased binding affinity to BS69, at approximately 60-fold less, compared to a wildtype control (Figure 3.7B). We wanted to ensure the decrease in activity was not due to decreased protein expression levels, so we performed a western blot on yeast cell lysate for each vector (Figure 3.7C). All the proteins, except for the M12 bait, were expressed as expected. G6PD was used as the loading control.

Figure 3.6. Using alanine scanning mutagenesis to determine the critical residues within the E1A minimal interacting region required to bind to BS69. A) Sequences of the E1A alanine scanning mutants with the corresponding designations. The PXLXP interaction motif is highlighted in red. The alanine substitutions in the sequences are also in red and bold. B) Yeast two-hybrid assay of E1A alanine scanning mutants, in the context of CR2, in bait and BS69 MYND in prey in W303-1A yeast. β-galactosidase activity in Miller units is shown on the Y-axis, while the alanine scanning mutants are indicated on the X-axis. The negative control (Neg) consisted of yeast transformed with wildtype E1A CR2 in bait and an empty prey vector, while wildtype (WT) the positive control consisted of yeast transformed with E1A CR2 in bait and BS69 MYND in prey. Results are shown as mean ± SD, n=3. Significance markers are assigned in comparison to the WT positive control (* $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test).

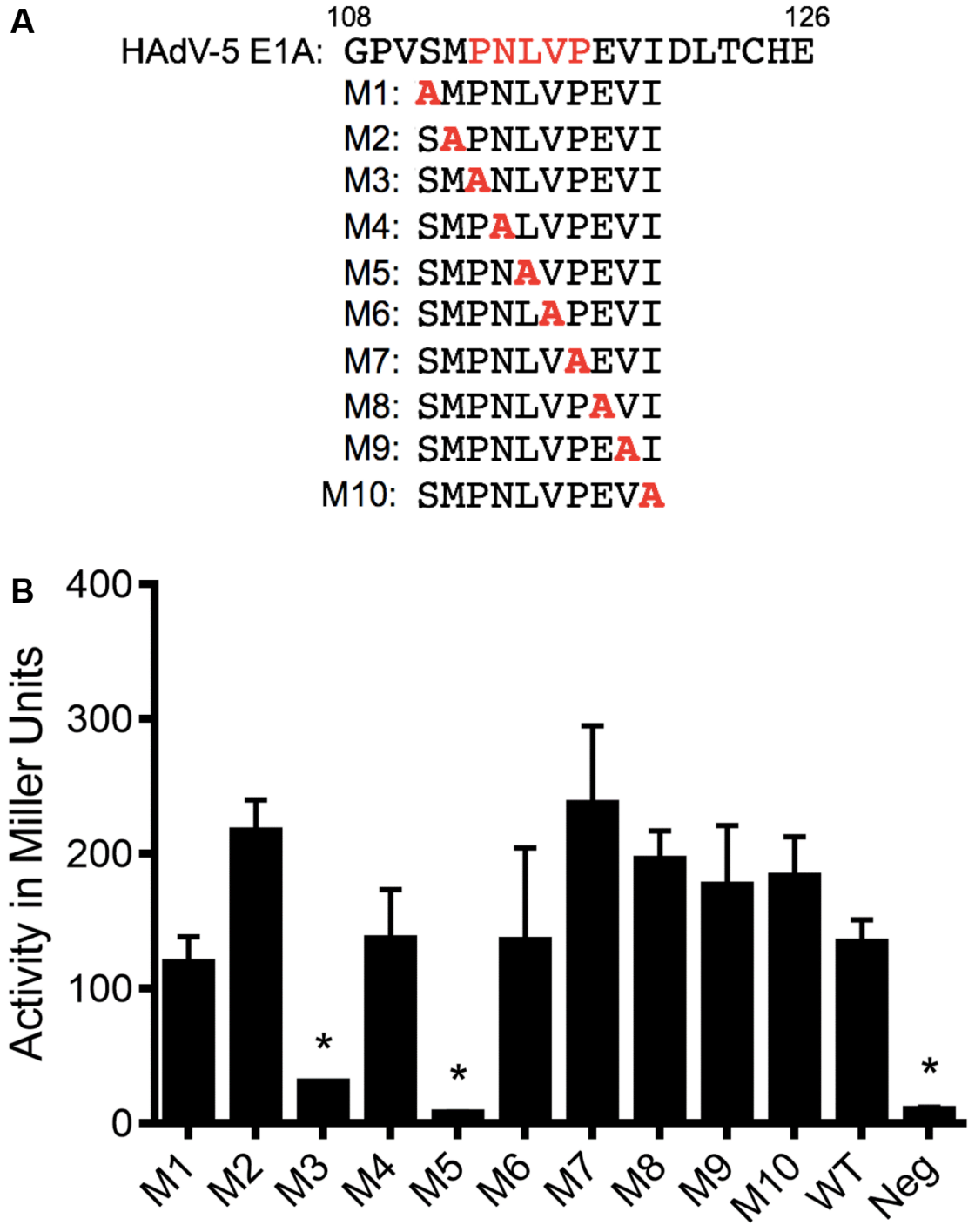
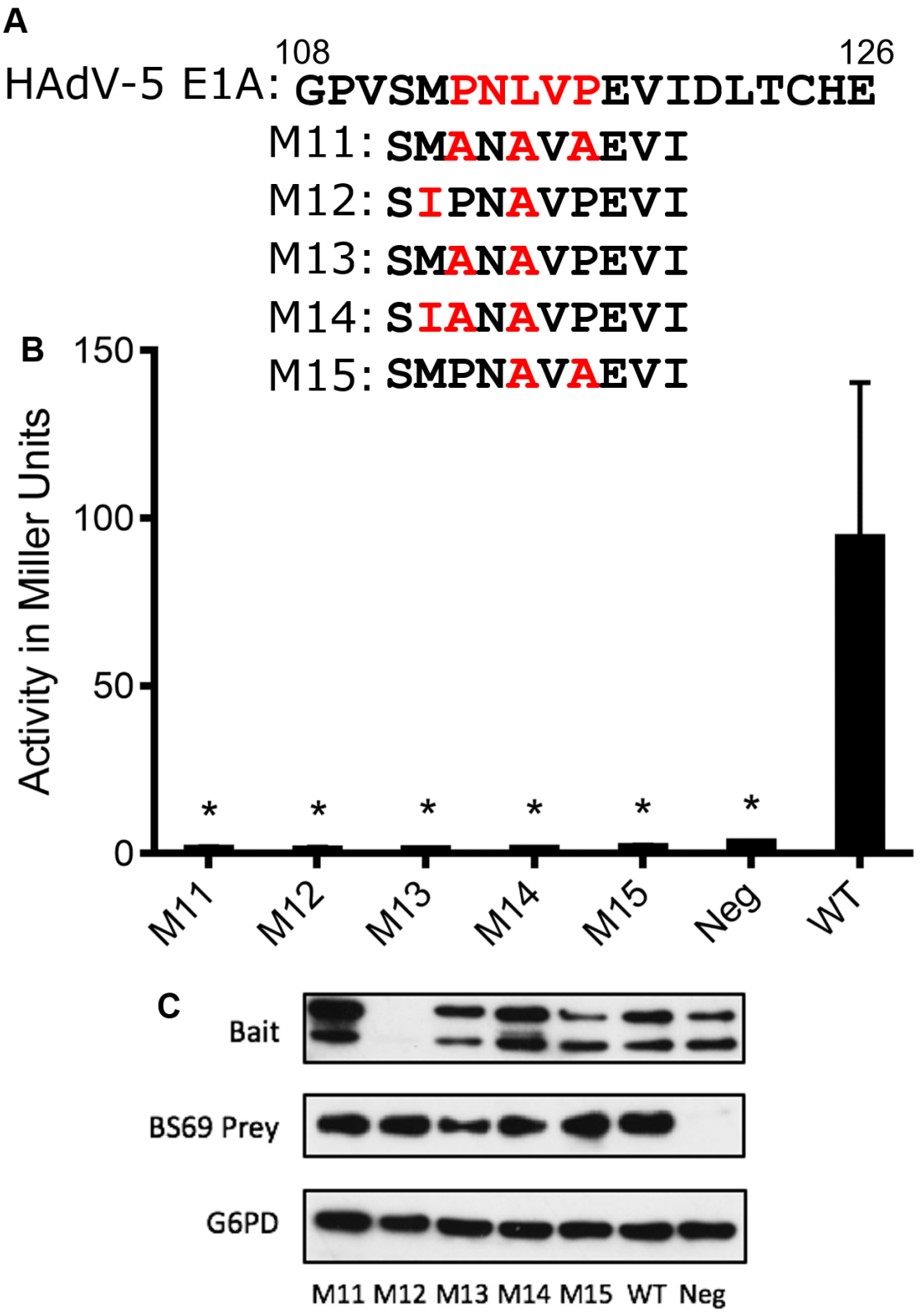

A
108
126
HAdV-5 E1A: GPVSMPNLVPEVIDLTCHE
M1: AMPNLVPEVI
M2: SAPNLVPEVI
M3: SMANLVPEVI
M4: SMPALVPEVI
M5: SMPNAVPEVI
M6: SMPNLAPEVI
M7: SMPNLVAEVI
M8: SMPNLVPAVI
M9: SMPNLVPEAI
M10: SMPNLVPEVA
B
Activity in Miller Units
400
300
200
100
0
*
*
*
M1
M2
M3
M4
M5
M6
M7
M8
M9
M10
WT
Neg

Figure 3.7. Introducing double and triple point mutations within E1A to thoroughly abrogate its ability to interact with BS69. A) Sequences of the E1A double and triple point mutants with the corresponding designations. The PXLXP interaction motif is highlighted in red. The double and triple point mutants in the sequences are also in red and bold. B) Yeast two-hybrid assay of E1A double and triple point mutants, in the context of CR2, in bait and BS69 MYND in prey in W303-1A yeast. β-galactosidase activity in Miller units is shown on the Y-axis, while the double and triple point mutants are indicated on the X-axis. The negative control (Neg) consisted of yeast transformed with E1A CR2 in bait and an empty prey vector, while the positive control (WT) consisted of yeast transformed with wildtype E1A CR2 in bait and BS69 MYND in prey. Results are shown as mean ± SD, n=3. Significance markers are assigned in comparison to the WT positive control (* $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test). C) Western blot of yeast cell lysate to verify protein expression. Bait and prey proteins were visualized using anti-LexA DBD and anti-HA antibodies respectively. G6PD was used as a loading control.


A
108
126
HAdV-5 E1A: GPVSMPNLVPEVIDLTCHE
M11: SMANAVAEVI
M12: SIPNAVPEVI
M13: SMANAVPEVI
M14: SIANAVPEVI
M15: SMPNAVAEVI
B
Activity in Miller Units
150
100
50
0
M11
M12
M13
M14
M15
Neg
WT
C
Bait
BS69 Prey
G6PD
M11 M12 M13 M14 M15 WT Neg

3.1.4 Purifying the MYND domain of BS69

With the goal of using recombinant BS69 to probe an array of peptides based on the E1A interaction sequence, we purified the MYND domain of BS69 using affinity chromatography. The MYND domain of BS69 was fused to GST or 6xHis epitope tags using the pGEX4T1 or pET28a backbone vectors respectively. The vectors were then transformed into BL-21 *Escherichia coli*. Selected colonies were expanded in LB liquid culture medium until OD_{600} reached 0.5, before treatment with IPTG to induce exogenous protein expression and further growth overnight at 16^{o}C. The cell lysates were then eluted through a glutathione or Nickel-NTA column for GST and 6xHis tagged proteins respectively. Samples were collected at various steps to measure the quality of protein purification. MYND-His in the pET28a backbone has a molecular weight of 25kDa, which corresponds to the prominent band in the elution lanes on the gel stained with Coomassie Blue (Figure 3.8A). Some samples were also analyzed using a western blot probed with an anti-BS69 antibody, and the bands in the elution lanes show a laddering effect with the molecular weights in multiples of 25kDa (Figure 3.8B). This indicates that the protein is forming polymers, possibly due to the shortage of zinc ions that are coordinated by cysteine and histidine residues within the MYND domain. Thus, we repeated the purification protocol of both GST and His tagged MYND, except we supplemented the growth media with 100μM ZnCl. MYND-GST has a molecular weight of 47kDa, and is the only detectable band in the elution lanes on the gel stained with Coomassie Blue (Figure 3.9). The MYND-His purification resulted in a similar blot to the previous attempt, in that the 25kDa band is still prominent on the Coomassie Blue gel. We have also completed preliminary protein purification of pRb-GST (residues 385-773) and UBC9-GST at three different post-IPTG incubation temperatures, and detected the protein using a western blot probed using an anti-GST antibody (Figure 3.10).

3.1.5 Mapping the E1A-BS69 Binding Site Using a Peptide Array

Peptide arrays are a powerful tool used for high-throughput epitope mapping, enzyme profiling, phosphorylation studies, and protein-protein interactions. These arrays provide

Figure 3.8. Coomassie Blue gel and Western blot showing His tagged MYND protein purification. 6xHis-tagged BS69 MYND was expressed in BL21 (DE3) RIL *Escherichia coli* and induced using 1mM IPTG. Samples were collected before adding IPTG (Pre-IPTG), after 24 hour IPTG induction (Post-IPTG 1), after -80^{o}C overnight freezing (Post-IPTG 2), after being passed through the Nickel-NTA column (Flow Through), and after 11 elution fractions. A) Coomassie Blue-stained gel of the various samples collected throughout the purification process. The ladder is labelled in kDa and shown on the leftmost lane. Expected molecular weight of 6xHis-tagged BS69 MYND is 27kDa. B) western blot showing elution fractions 1, 4, 7 and Post-IPTG 1. The ladder is labelled on the left in kDa and the membrane was probed with an anti-BS69 antibody.

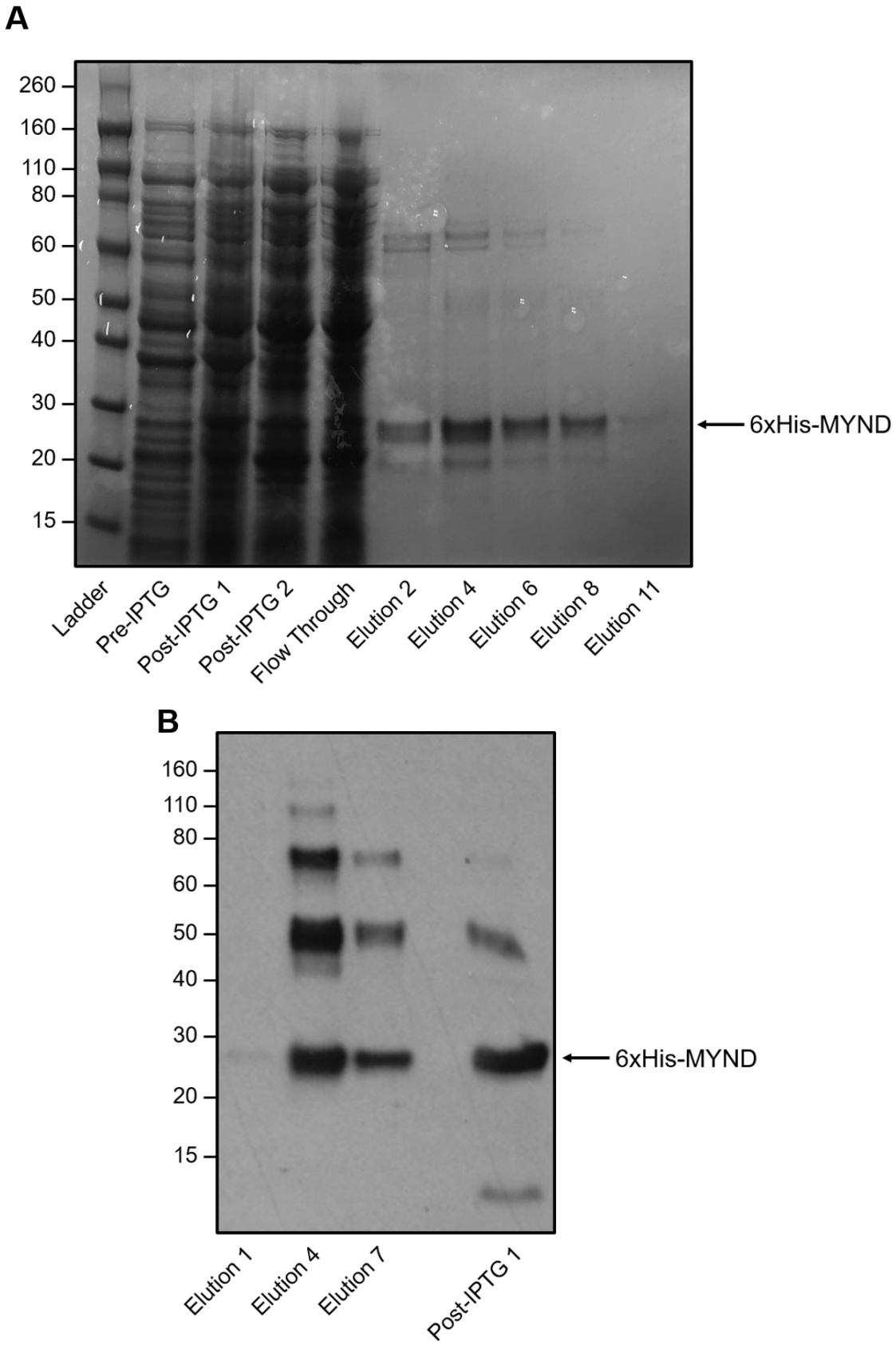

A
260
160
110
80
60
50
40
30
20
15
6xHis-MYND
Ladder
Pre-IPTG
Post-IPTG 1
Post-IPTG 2
Flow Through
Elution 2
Elution 4
Elution 6
Elution 8
Elution 11
B
160
110
80
60
50
40
30
20
15
6xHis-MYND
Elution 1
Elution 4
Elution 7
Post-IPTG 1

Figure 3.9. Coomassie Blue gel showing His and GST tagged MYND protein purification. Coomassie Blue stained gel of the various samples collected throughout the protein purification process. GST and 6xHis-tagged BS69 MYND were expressed in BL21 (DE3) RIL *Escherichia coli* grown in 100μM ZnCl and induced using 1mM IPTG. Samples were collected before adding IPTG (Pre-IPTG), after 24 hour IPTG induction (Post-IPTG), and after 10 elution fractions. The ladder is labelled in kDa and shown on the leftmost lane. Expected molecular weight of GST and 6xHis-tagged BS69 MYND are 47kDa and 27kDa, respectively.

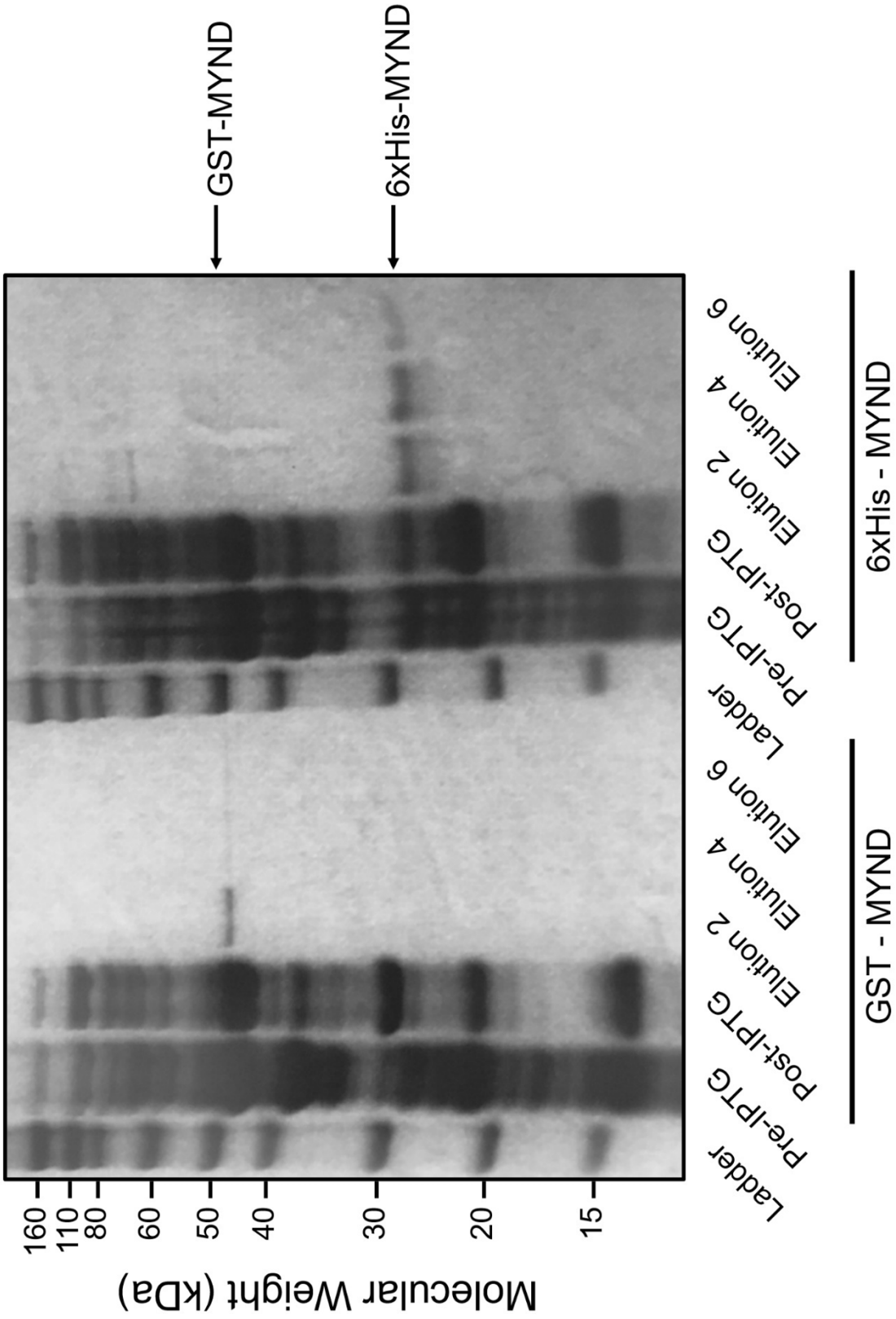
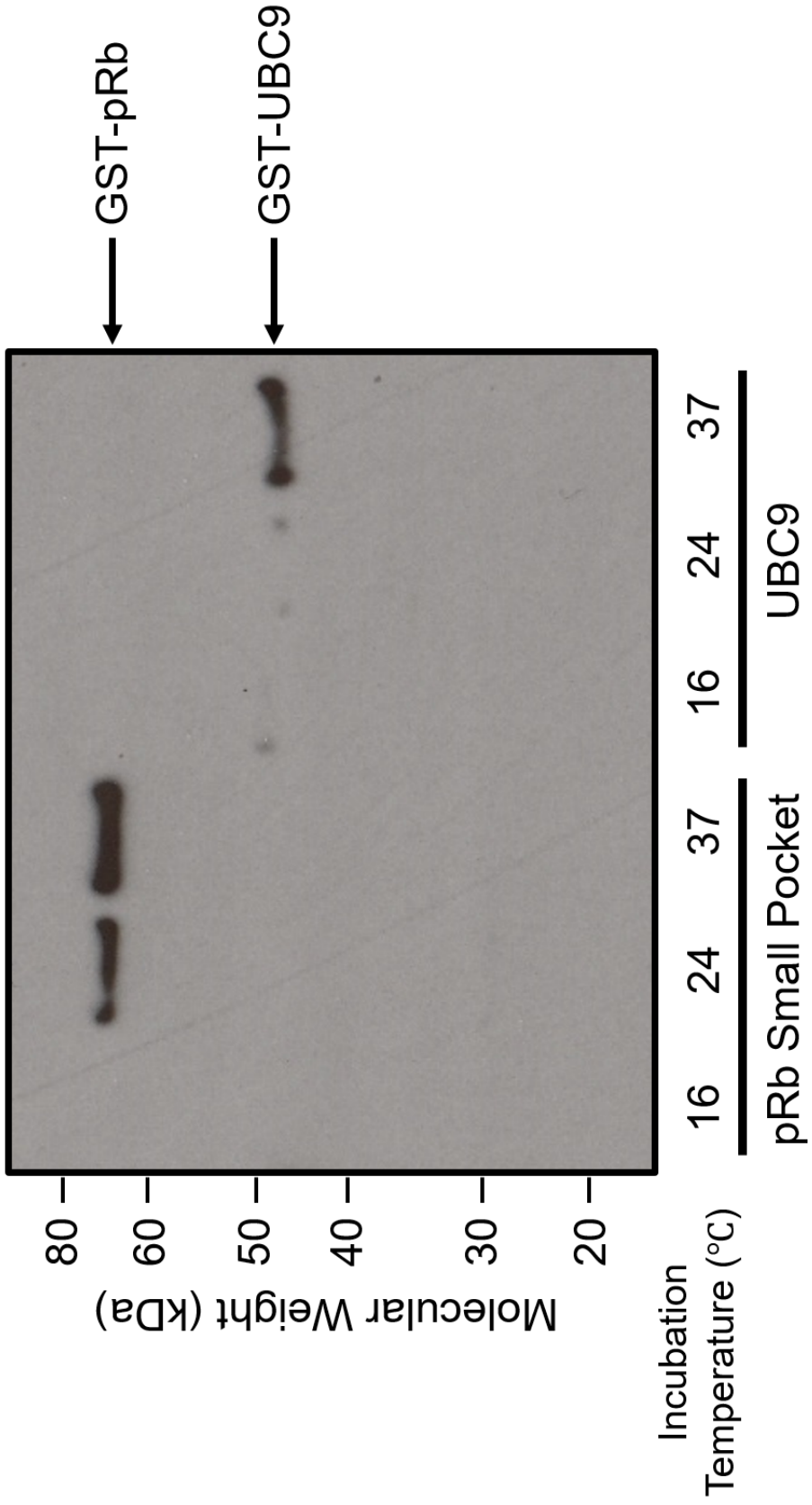

GST-MYND
6xHis-MYND
Molecular Weight (kDa)
160
110
80
60
50
40
30
20
15
Ladder
Pre-IPTG
Post-IPTG
Elution 2
Elution 4
Elution 6
Ladder
Pre-IPTG
Post-IPTG
Elution 2
Elution 4
Elution 6
GST - MYND
6xHis - MYND

Figure 3.10. Western blot showing protein purification of GST tagged pRb and UBC9. Western blot showing purification of GST-tagged pRb and UBC9. GST-tagged pRb and UBC9 were expressed in BL21 (DE3) RIL *Escherichia coli* and induced using 1mM IPTG for 4, 8, and 16 hours at 37^{o}C, 24^{o}C, and 16^{o}C, respectively. The ladder is labelled on the left in kDa and the membrane and the membrane was probed with an anti-GST antibody.


GST-pRb
GST-UBC9
80
60
50
40
30
20
Molecular Weight (kDa)
Incubation Temperature (°C)
16
24
37
16
24
37
pRb Small Pocket
UBC9

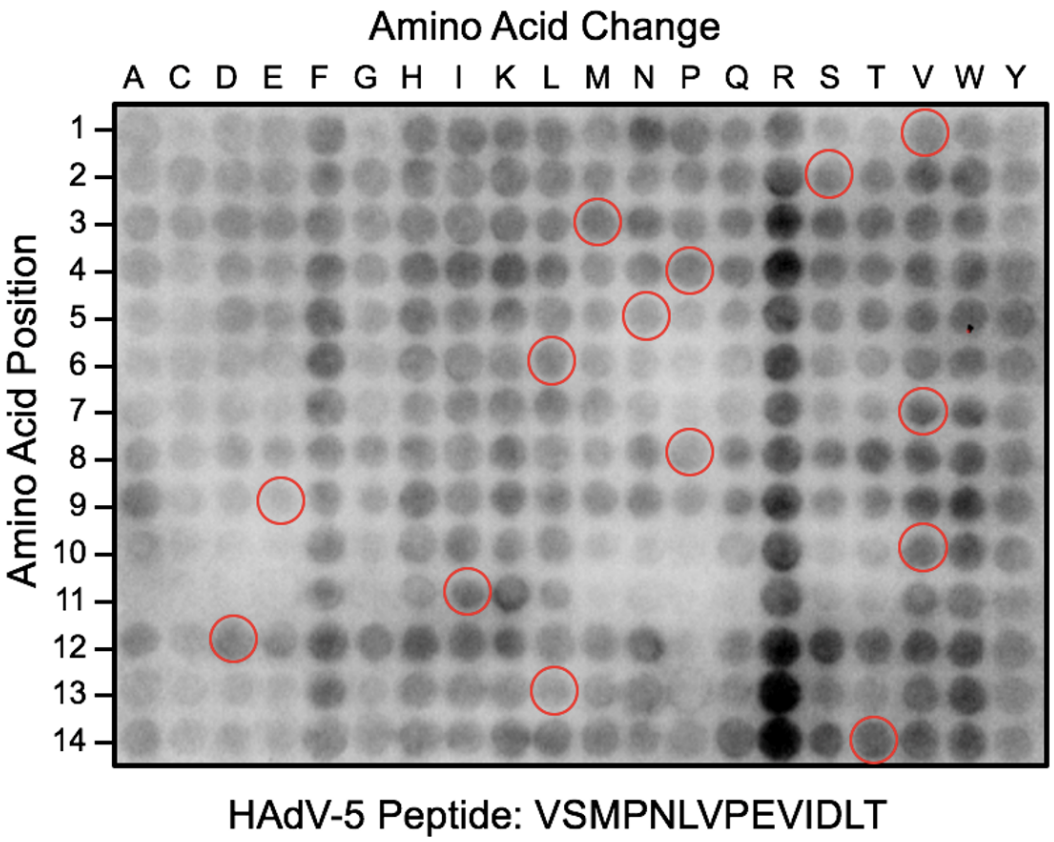
quick results and are cost-efficient to run. Using information from our minimal interacting region experiments, we designed a peptide array to characterize the protein-protein interaction between E1A and BS69. Because the amino acids at every position in the interaction motif under investigation were mutated to every other residue, this experiment allows us to address a major shortcoming of the alanine scanning panel we previously completed, in which every residue is replaced only with alanine (Figure 3.6).

We used positional scanning mutagenesis on 14-residue E1A sequences from HAdV-5 and HAdV-12 (wildtype HAdV-5 E1A sequence: VSMPNLVPEVIDLT and wildtype HAdV-12 E1A sequence: ECMPQLHPEDMDLL) to generate a peptide library. These peptides were synthesized directly onto a Whatman filter paper membrane using Fmoc solid-phase synthesis. We then blocked the membrane and probed it with purified MYND-His followed by probing with HRP-conjugated anti-His antibody (Figure 3.11). The blot showed high background with intense signals in the lane where the amino acids were sequentially mutated into arginine. Most of the spots corresponding to the wildtype sequence (circled in red) did not show strong signal strength. We stripped the membrane and re-probed using purified MYND-His followed by anti-BS69 antibody and a secondary HRP-conjugated anti-Rabbit IgG antibody (Figure 3.12). The results were similar to the previous attempt, as there was high background with intense signals in the arginine lane and weak signal strength for the wildtype sequences. Finally, we stripped the membranes and re-probed with the HRP-conjugated anti-His antibody without prior incubation of the purified MYND-His protein (Figure 3.13). We found that the high background signal remained on many of the spots. This protocol should be optimized with more stringent blocking steps, higher purity protein, and a more specific antibody for future experiments.

3.1.6 Verifying the E1A-BS69 Interaction in a Mammalian Cell Culture System

The previous experiments investigating the E1A-BS69 interaction in this project were completed either using purified proteins or in yeast. These model systems have some definite advantages, such as being able to measure direct protein-protein interactions or

Figure 3.11. Using an anti-6xHis antibody to probe a peptide array showing the interaction between the BS69 MYND domain and short peptide sequences from HAdV-5 and HAdV-12 E1A. Peptide arrays were provided by the Shawn Li lab and synthesized using Fmoc solid-phase peptide synthesis on an amino-functionalized cellular membrane. Peptides underwent positional scanning mutagenesis, with the corresponding wildtype sequence below the membrane. Each row denotes a position on the peptide that is being targeted, with the position labelled left of the membrane. Each column represents the amino acid, labelled on the top, that the target residue is mutated to. Membranes were blocked overnight with shaking in 3% BSA, then incubated with 1μM of purified 6xHis-tagged BS69 MYND overnight. The membranes were probed using an HRP-conjugated anti-6xHis tag antibody. Peptide spots circled in red denote the wildtype sequence.


Amino Acid Change
A C D E F G H I K L M N P Q R S T V W Y
Amino Acid Position
1 2 3 4 5 6 7 8 9 10 11 12 13 14
HAdV-5 Peptide: VSMPNLVPEVIDLT

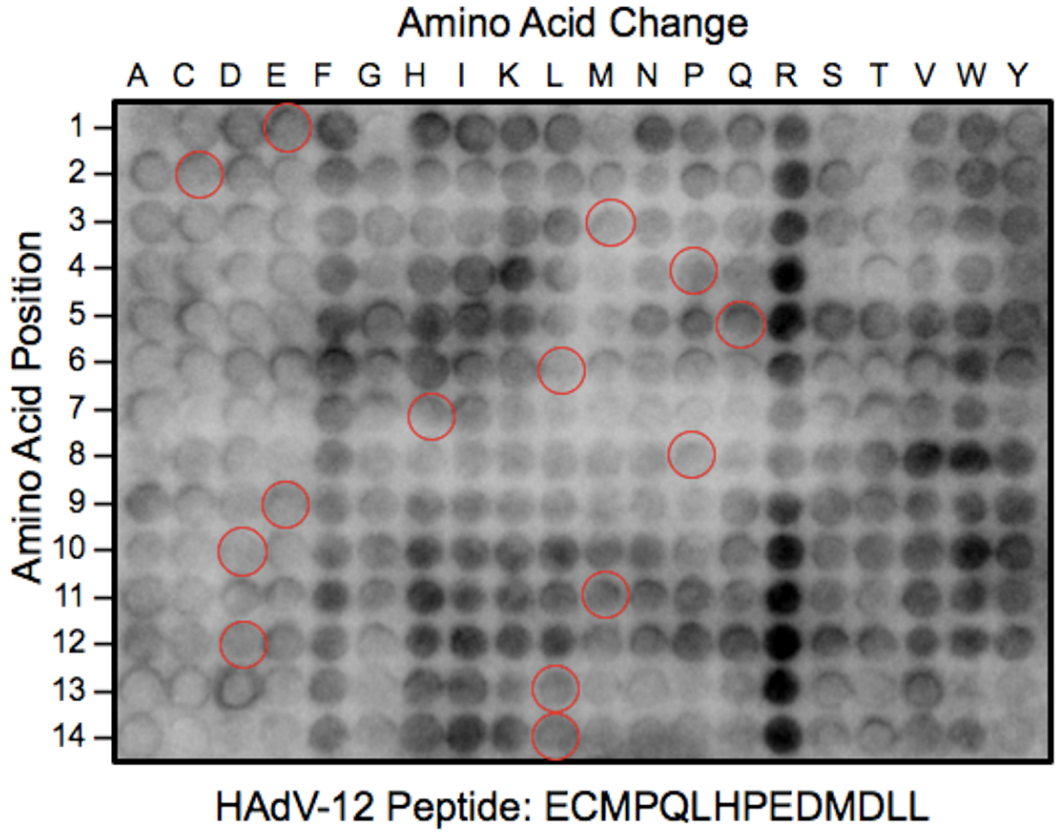
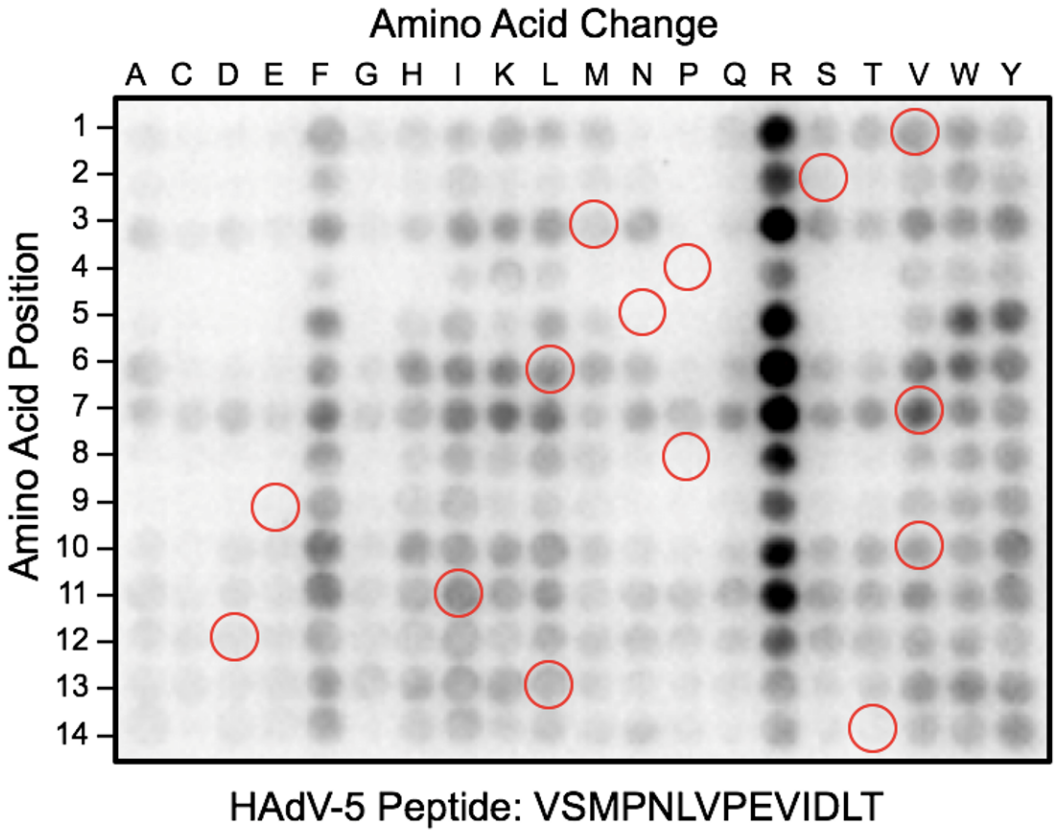

Amino Acid Change
A C D E F G H I K L M N P Q R S T V W Y
Amino Acid Position
1 2 3 4 5 6 7 8 9 10 11 12 13 14
HAdV-12 Peptide: ECMPQLHPEDMDLL

Figure 3.12. Using an anti-BS69 antibody to probe a peptide array showing the interaction between the BS69 MYND domain and short peptide sequences from HAdV-5 and HAdV-12 E1A. Peptide arrays were provided by the Shawn Li lab and synthesized using Fmoc solid-phase peptide synthesis on an amino-functionalized cellular membrane. Peptides underwent positional scanning mutagenesis, with the corresponding wildtype sequence below the membrane. Each row denotes a position on the peptide that is being targeted, with the position labelled left of the membrane. Each column represents the amino acid, labelled on the top, that the target residue is mutated to. Membranes were blocked overnight with shaking in 3% BSA, then incubated with 1μM of purified 6xHis-tagged BS69 MYND overnight. The membranes were probed using an anti-BS69 antibody followed by an HRP-tagged anti-rabbit IgG secondary. Peptide spots circled in red denote the wildtype sequence.


Amino Acid Change
A C D E F G H I K L M N P Q R S T V W Y
Amino Acid Position
1 2 3 4 5 6 7 8 9 10 11 12 13 14
HAdV-5 Peptide: VSMPNLVPEVIDLT

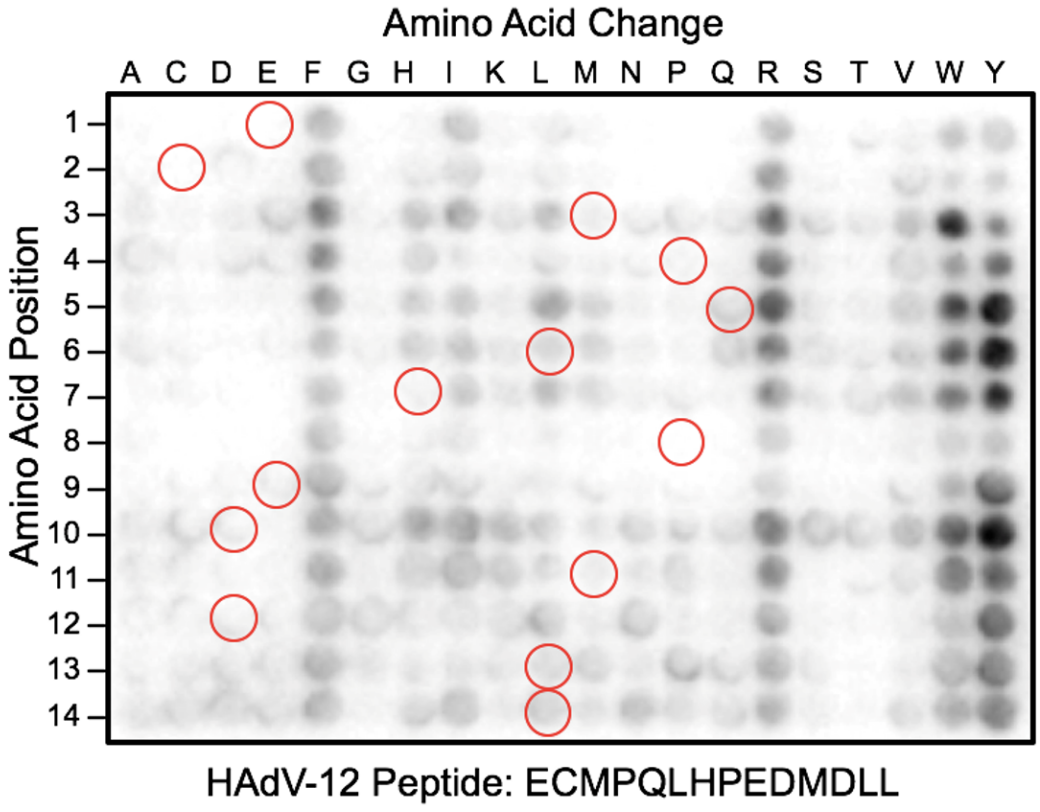
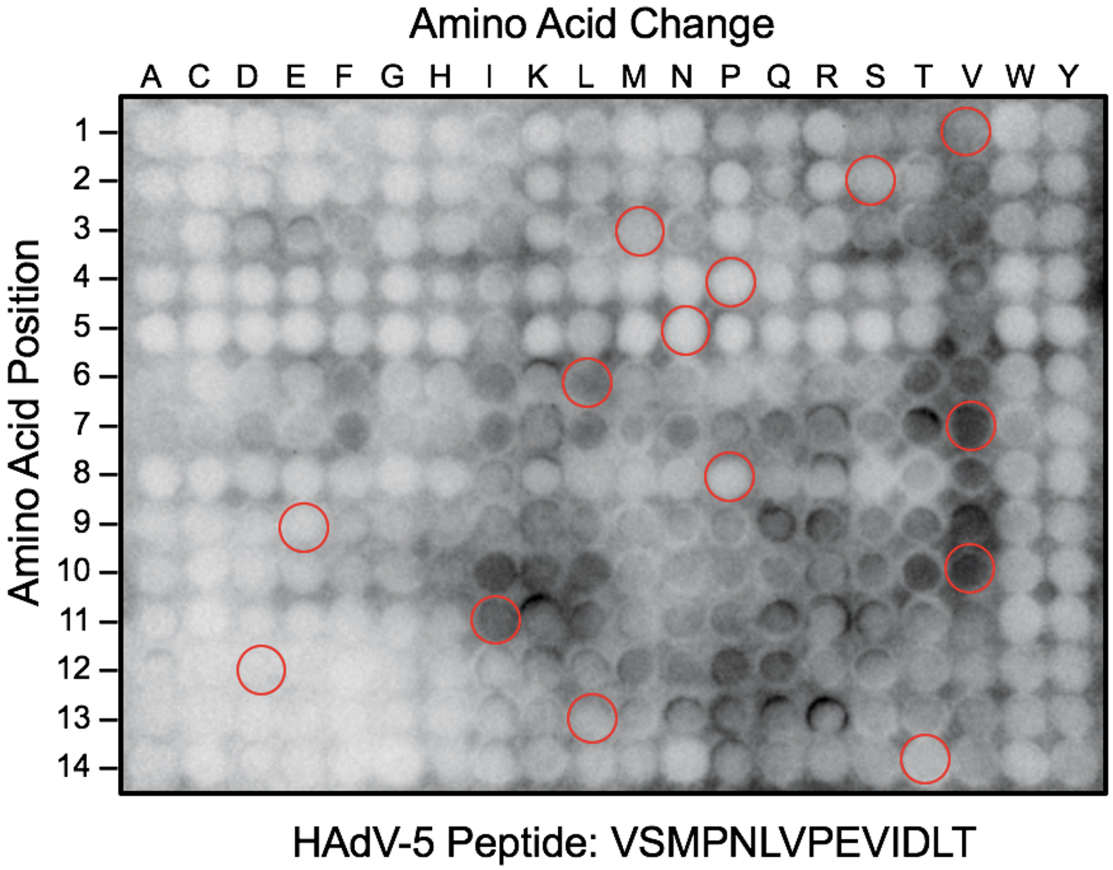

Amino Acid Change
A C D E F G H I K L M N P Q R S T V W Y
Amino Acid Position
1 2 3 4 5 6 7 8 9 10 11 12 13 14
HAdV-12 Peptide: ECMPQLHPEDMDLL

Figure 3.13. Using an anti-His antibody to probe the stripped HAdV-5 E1A peptide array. Peptide arrays were provided by the Shawn Li lab and synthesized using Fmoc solid-phase peptide synthesis on an amino-functionalized cellular membrane. Peptides underwent positional scanning mutagenesis, with the corresponding wildtype sequence below the membrane. Each row denotes a position on the peptide that is being targeted, with the position labelled left of the membrane. Each column represents the amino acid, labelled on the top, that the target residue is mutated to. The membranes were stripped using a sonication bath in a solution containing urea followed by an acetic acid stripping solution. The peptide array was then blocked overnight shaking in 3% BSA. Without incubating with purified protein, the membrane was probed using an HRP-conjugated anti-6xHis tag antibody. Peptide spots circled in red denote the wildtype sequence.


Amino Acid Change
A C D E F G H I K L M N P Q R S T V W Y
Amino Acid Position
1
2
3
4
5
6
7
8
9
10
11
12
13
14
HAdV-5 Peptide: VSMPNLVPEVIDLT

allowing us to quickly screen multiple constructs. However, using these artificial systems may compromise biological relevance as the background by which this interaction is studied is different from that of a mammalian cellular environment. Thus, we set out to verify the E1A-BS69 interaction in a mammalian system by utilizing co-IP experiments in HT1080 human fibrosarcoma cells.

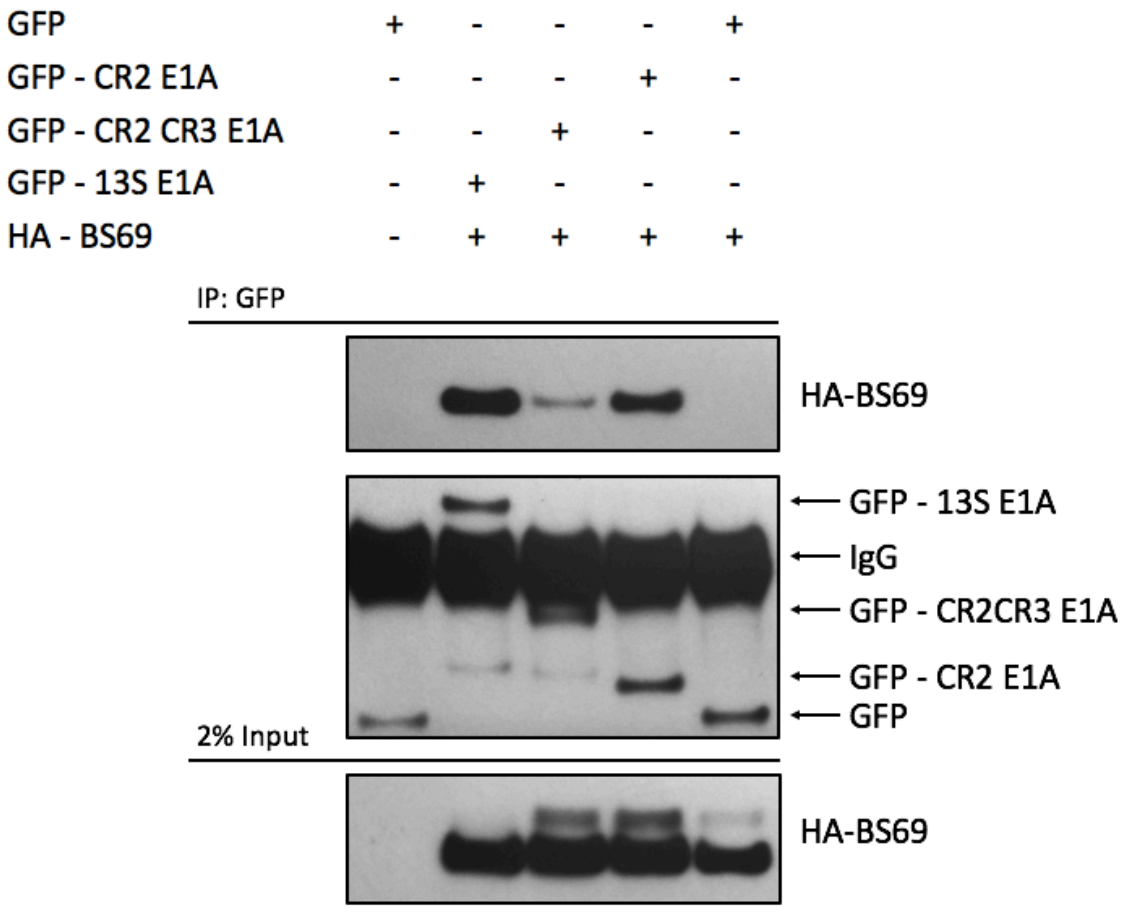
We co-transfected HT1080 cells with equal amounts of HA-tagged BS69 and GFP-tagged E1A constructs. These constructs included CR2 (residues 93-139), CR2 CR3 (residues 93-191), and full length E1A (13S, residues 1-289). Cells transfected with an empty GFP vector with or without HA-tagged BS69 were used as negative controls. Cells were lysed 24 hours post transfection and subject to co-IP using protein A-sepharose beads and an anti-GFP antibody. The samples were then run on a western blot and probed with an anti-HA antibody to determine which E1A constructs could pull down BS69 (Figure 3.14). All tested E1A constructs were able to pull down BS69; 13S E1A had the highest affinity, followed sequentially by CR2 E1A and CR2 CR3 E1A (Figure 3.14). The faint band showing weak interaction between BS69 and E1A CR2 CR3 may be due to weak expression or improper folding within CR3 of this E1A construct. BS69 was not pulled down by the GFP negative control.

3.2 Analysis of BS69 Mediated Changes on E1A Induced Transactivation

3.2.1 Effects of Increasing BS69 Concentration on E1A Mediated Transactivation

Previous studies in the Bernards lab have shown that BS69 strongly inhibits E1A mediated transactivation (Hateboer et al., 1995). In these experiments, the authors used luciferase assays to measure transcriptional changes of a reporter vector in cells transfected with GAL4-E1A and increasing concentrations of BS69-HA (Hateboer et al., 1995). Other studies have shown that the reverse is also true: E1A is able to antagonize BS69 mediated repression of cellular genes in a manner dependent on the fidelity of the E1A PXLXP motif (Ansieau & Leutz, 2002). Thus, we set out to combine various aspects of all these studies

Figure 3.14. Co-immunoprecipitation of E1A constructs and BS69 in HT1080 cells. HT1080 cells were co-transfected with 4μg each of HA-tagged BS69 MYND and GFP-tagged E1A constructs. Negative controls consisted of cells transfected with 4μg of GFP vector with and without being cotransfected with 4μg of HA-tagged BS69 MYND. Cell lysates were collected 24 hours post transfection and subject to immunoprecipitation using an anti-GFP antibody. Samples were subsequently run on a western blot and probed with anti-GFP and anti-HA antibodies.

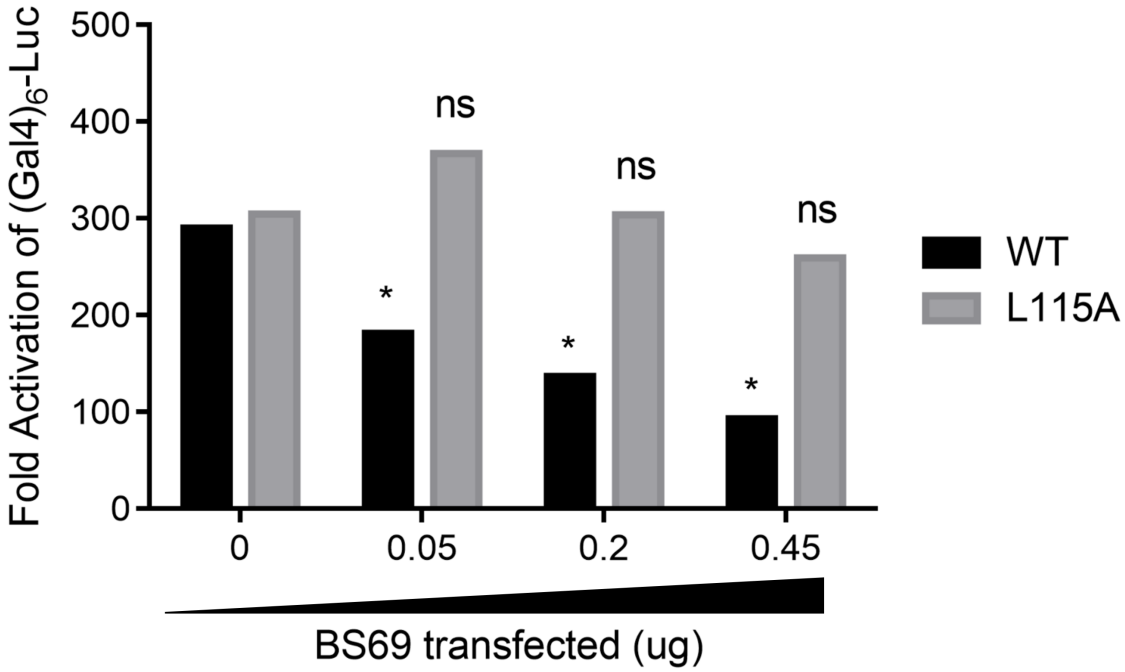

GFP + - - - +
GFP - CR2 E1A - - - + -
GFP - CR2 CR3 E1A - - + - -
GFP - 13S E1A - + - - -
HA - BS69 - + + + +
IP: GFP
HA-BS69
GFP - 13S E1A
IgG
GFP - CR2CR3 E1A
GFP - CR2 E1A
GFP
2% Input
HA-BS69

to further explore the effects of BS69 on E1A transactivation. Using luciferase assays, we verified that BS69 functions as a dose dependent repressor of E1A transactivation, and determined that the PXLXP motif is necessary for this observation (Figure 3.15).

We co-transfected HT1080 cells with a GAL4-responsive luciferase reporter plasmid, which contains six GAL4 binding sites upstream of the luciferase reporter gene; a plasmid expressing 13S wildtype E1A, or 13S L115A E1A in the GAL4-DBD backbone; increasing concentrations of BS69 expressed in the pcDNA3-HA backbone; a pSV-β-galactosidase vector, used to measure β-galactosidase activity as an internal transcription control; and an empty pcDNA3-HA plasmid as a "stuffer" as required to ensure that 2μg of total DNA was transfected per sample (Figure 3.15). β-galactosidase activity was consistent between all samples (not shown). 0μg, 0.05μg, 0.2μg, and 0.45μg of BS69 constructs were transfected per well. The negative controls consisted of HT1080 cells transfected with the reporter plasmid, an empty GAL4-DBD plasmid, and the empty pcDNA3-HA "stuffer" plasmid (not shown). The luciferase activity was then measured from the cell lysate, which was obtained 24 hours post-transfection. Results were normalized to total protein concentration as determined by Bradford assays, then reported as fold activation over the empty GAL4-DBD negative control.

In the series of experiments with wildtype E1A, increasing concentrations of BS69 resulted in a stepwise and significant decrease in luciferase activity (Figure 3.15). Compared to the luciferase activity of cells not transfected with BS69, cells transfected with 0.05μg, 0.2μg, and 0.45μg of BS69 showed a 37.4%, 52.9%, and 67.9% decrease in luciferase activity respectively. On the other hand, increasing the concentration of BS69 did not significantly change the luciferase activity in samples transfected with the L115A E1A construct, which is unable to bind BS69 due to the altered PXLXP motif. When comparing cells transfected with equal amounts of BS69, luciferase activity was also higher in cells transfected with L115A E1A than wildtype E1A.

Figure 3.15. BS69 represses E1A mediated transactivation through the PXLXP motif. Luciferase assay showing BS69-mediated repression of E1A transactivation. Luciferase activity, shown as fold activation over the negative control is labelled on the Y-axis, while the amount of BS69 transfected is shown on the X-axis. Black bars represent samples co-transfected with wildtype E1A, while grey bars represent samples co-transfected with L115A E1A. HT1080 cells were co-transfected with 1µg of pGL2-$(UAS)_6$-Luc reporter, 0.5µg of pSV-β-galactosidase, 0.05µg pM 13S Ad5 E1A or pM 13S E1A L115A, increasing concentrations of HA BS69, and pcDNA3 HA to ensure equal loading across all samples. Samples encoding an empty pM vector in place of E1A was used as the negative control (not shown). Cell lysates were collected 24 hours post transfection and the luciferase activity was measured. Results were normalized by protein concentration and reported as fold activity compared to the negative control. Results are shown as mean, n=2. Significance markers are assigned in comparison to the samples not transfected with a BS69 vector (ns = not significant; * $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test).


Fold Activation of (Gal4)6-Luc
500
400
300
200
100
0
ns
ns
ns
*
*
*
0
0.05
0.2
0.45
BS69 transfected (ug)
WT
L115A

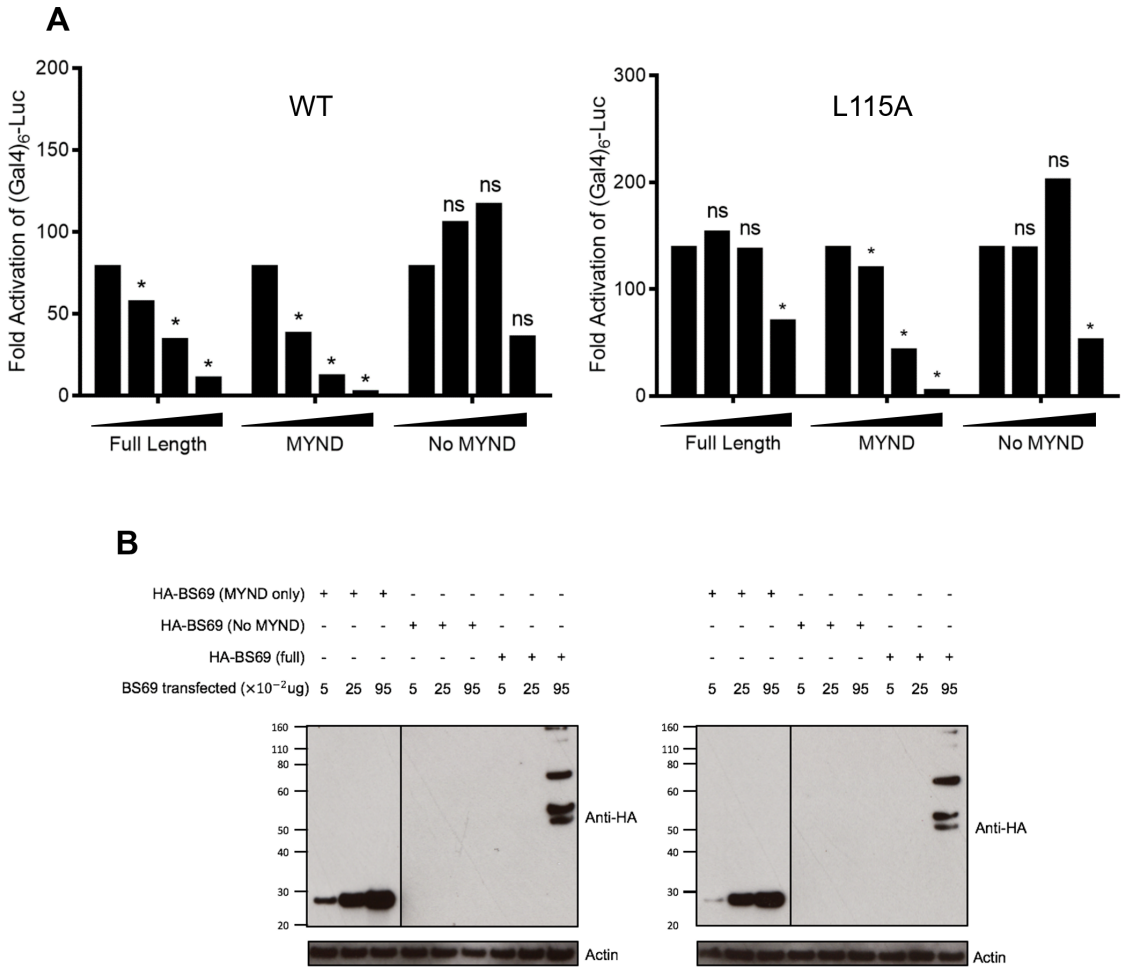
3.2.2 Determining the Region of BS69 Responsible for Repression of E1A Mediated Transactivation

BS69 is a 602-amino acid protein consisting of four known domains. The PHD, Bromo, and PWWP domains are located sequentially in the N-terminal half of the protein. These three domains are all involved in mediating the interaction between BS69 and histone H3.3K36me3, with the PWWP domain being the most crucial in this interaction (Wang et al., 2014; Wen et al., 2014). The MYND domain is a zinc finger near the C-terminus of the protein, and this domain is used to form interactions with other proteins including E1A. Thus, we set out to determine the domains of BS69 required to repress E1A mediated transactivation.

We co-transfected HT1080 cells with a GAL4-responsive luciferase reporter plasmid; a plasmid expressing 13S wildtype E1A or 13S L115A E1A in a GAL4-DBD backbone; increasing concentrations of full length (residues 47-602), MYND only (residues 427-602), or NO MYND (residues 47-426) constructs of BS69 in the pcDNA3-HA backbone; and the pcDNA3-HA "stuffer" plasmid as necessary to ensure that 2µg of total DNA was transfected per sample (Figure 3.16). The pSV-β-galactosidase vector was not used in these subsequent experiments in order to allow for increased concentrations of transfected BS69. 0µg, 0.05µg, 0.25µg, and 0.95µg of BS69 constructs were transfected per well. The negative controls consisted of HT1080 cells transfected with the reporter plasmid, an empty GAL4-DBD plasmid, and the empty pcDNA3-HA "stuffer" plasmid (not shown). The luciferase activity was then measured from the cell lysate, which was obtained 24 hours post-transfection. Results are normalized to total protein concentration as determined by Bradford assays, then reported as fold activation over the empty GAL4-DBD negative control. Significance markers are assigned in comparison to the activity of samples transfected with no BS69 within the respective group (Figure 3.16).

The samples transfected with 0.95µg of BS69, regardless of the E1A construct being evaluated, all resulted in less luciferase activity than the other samples in their respective groups (Figure 3.16A). This could be a non-specific artefact caused by protein overexpression. Cells transfected with wildtype E1A showed a stepwise decline in luciferase activity when increasing concentrations of BS69, either full length or MYND,

Figure 3.16. The MYND domain of BS69 is necessary and sufficient in repressing E1A mediated transactivation. A) Luciferase assay showing BS69-mediated repression of E1A transactivation. Luciferase activity, shown as fold activation over the negative control is labelled on the Y-axis, while the relative amount and construct of BS69 transfected is shown on the X-axis. HT1080 cells were co-transfected with 1μg of pGL2-(UAS)6-Luc reporter, 0.05μg pM 13S Ad5 E1A or pM 13S E1A L115A, increasing concentrations of HA BS69, and pcDNA3 HA to ensure equal loading across all samples. Samples encoding an empty pM vector in place of E1A was used as the negative control (not shown). Cell lysates were collected 24 hours post transfection and the luciferase activity was measured. Results were normalized by protein concentration and reported as fold activity compared to the negative control. Results are shown as mean, n=2. Significance markers are assigned in comparison to the samples not transfected with a BS69 vector (ns = not significant; * $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test). B) Western blot showing expression of BS69 constructs used in the luciferase assays. Cell lysates from the luciferase experiments were run on a western blot and probed using an anti-HA antibody to visualize the BS69 constructs, and anti-actin for the loading control.


A
WT
L115A
Fold Activation of (Gal4)6-Luc
200
150
100
50
0
300
200
100
0
ns
*
Full Length
MYND
No MYND
B
HA-BS69 (MYND only)
HA-BS69 (No MYND)
HA-BS69 (full)
BS69 transfected (×10−2ug)
5 25 95 5 25 95 5 25 95
160
110
80
60
50
40
30
20
Anti-HA
Actin

were introduced. Increasing concentrations of full length BS69 resulted in a 27.3%, 56.7%, and 85.5% decrease in luciferase activity at 0.05µg, 0.25µg, and 0.95µg BS69 transfections respectively compared to cells not transfected with BS69, while increasing concentrations of MYND resulted in a 51.8%, 84.9%, and 97.2% decrease in luciferase activity at 0.05µg, 0.25µg, and 0.95µg BS69 transfections respectively compared to cells not transfected with BS69. This change in activity was not observed with increasing concentrations of the No MYND BS69 construct, except for the sample transfected with 0.95µg of No MYND BS69. This sample showed a 54.6% decrease in activity compared to cells not transfected with BS69, albeit without statistical significance. Again, this result could be an artefact caused by the high levels of protein expression from this sample. These results suggest that the MYND domain is the primary mediator of repression of E1A transactivation.

Unexpectedly, cells transfected with the L115A E1A mutant, which does not bind BS69 efficiently, also showed a stepwise decline in luciferase activity with increasing concentrations of BS69 MYND. This resulted in a 13.8%, 69.1%, and 96.6% decrease in activity at 0.05µg, 0.25µg, and 0.95µg MYND transfections respectively compared to cells not transfected with BS69. Increasing concentrations of full length BS69 did not influence luciferase activity, except at the highest amount, 0.95µg, where the activity was decreased significantly by 49.5%. Increasing concentrations of No MYND BS69 had a similar effect on L115A E1A as it did on wildtype E1A, where the highest concentration sample showed a significant decrease of 62.4% in activity.

Next, we wanted to see if the change in plasmid transfection concentration caused a stepwise increase in protein levels, so we ran the cell lysate on a western blot (Figure 3.16B). BS69 MYND expressed very strongly, and showed the expected increase in protein levels. We could only detect full length BS69 on the western blot at the highest transfected concentration, and we were unable to detect the No MYND construct. The dark line separating the blot denotes a difference in exposure time of the chemiluminescent film, whereby the MYND domain was detectable within 5 seconds of exposure, while full length BS69 took approximately 15 minutes. This dramatic difference in expression levels between the full-length MYND construct and the MYND only construct may explain why residual repression activity was detected when MYND only construct was introduced into

samples transfected with the L115A E1A mutant. Actin was used as the loading control in these western blots (Figure 3.16B)

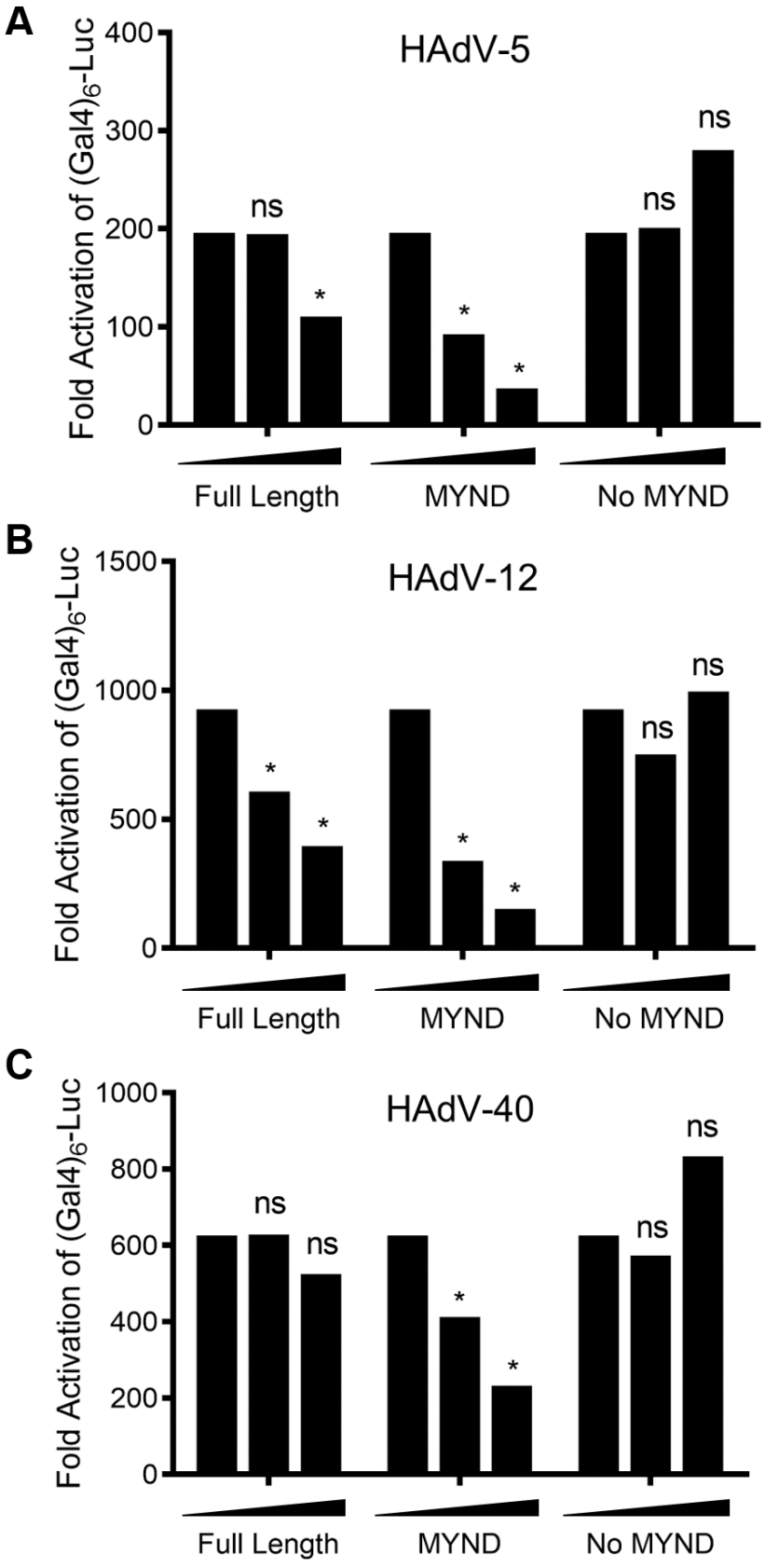
3.2.3 Establishing the Conservation of the BS69 Mediated Repression of E1A Transactivation in HAdV-5, 12, and 40

We have previously determined that adenovirus types 5, 9, and 12 from species C, D and A respectively are able to interact with BS69, while adenovirus types 3, 4, and 40 from species B, E, and F respectively could not. This observation may be due to the presence of the PXLXP motif in the E1A from adenovirus types able to interact with BS69, and the lack thereof on E1A from adenovirus types that are unable to interact with BS69. Thus, we set out to determine if the ability of E1A from a specific adenovirus species to bind to BS69 is related to its sensitivity to BS69 mediated repression of E1A transactivation.

We used a very similar protocol to the previous set of experiments, except we did not transfect any of these samples with 0.95μg of BS69. We co-transfected HT1080 cells with a GAL4-responsive luciferase reporter plasmid; a plasmid expressing 13S E1A from adenovirus types 5, 12, and 40, from species C, A, and F respectively, in a GAL4-DBD backbone; increasing concentrations of full length, MYND only, or NO MYND constructs of BS69 in the pcDNA3-HA backbone; and the pcDNA3-HA "stuffer" plasmid as necessary to ensure that 2μg of total DNA was transfected per sample (Figure 3.17). 0μg, 0.05μg, and 0.25μg of BS69 constructs were transfected per well. The negative controls consisted of HT1080 cells transfected with the reporter plasmid, an empty GAL4-DBD plasmid, and the empty pcDNA3-HA "stuffer" plasmid (not shown). The luciferase activity was then measured from the cell lysate, which was obtained 24 hours post-transfection. Results are normalized to total protein concentration as determined by Bradford assays, then reported as fold activation over the empty GAL4-DBD negative control. Significance markers are assigned in comparison to the activity of samples transfected with no BS69 within the respective group (Figure 3.17).

Transfection of increasing concentrations of No MYND did not cause significant changes in luciferase activity in cells transfected with E1A from all three tested adenovirus types

Figure 3.17. Sensitivity to BS69 mediated repression of E1A transactivation is proportional to the relative binding affinity of the two proteins between different adenovirus species. Luciferase assay showing BS69-mediated repression of E1A transactivation. Luciferase activity, shown as fold activation over the negative control is labelled on the Y-axis, while the relative amount and construct of BS69 transfected is shown on the X-axis. HT1080 cells were co-transfected with 1μg of pGL2-(UAS)6-Luc reporter; 0.05μg pM 13S HAdV-5 E1A (A), pM 13S HAdV-12 E1A (B), or pM 13S HAdV-40 E1A (C); increasing concentrations of HA BS69; and pcDNA3 HA to ensure equal loading across all samples. Samples encoding an empty pM vector in place of E1A was used as the negative control (not shown). Cell lysates were collected 24 hours post transfection and the luciferase activity was measured. Results were normalized by protein concentration and reported as fold activity compared to the negative control. Results are shown as mean, n=2. Significance markers are assigned in comparison to the samples not transfected with a BS69 vector (ns = not significant; * $p \leq 0.05$; one-way ANOVA with Dunnett's multiple comparisons test).


A
HAdV-5
Fold Activation of (Gal4)6-Luc
400
300
200
100
0
ns
ns
ns
*
*
*
Full Length
MYND
No MYND
B
HAdV-12
Fold Activation of (Gal4)6-Luc
1500
1000
500
0
*
*
*
*
ns
ns
Full Length
MYND
No MYND
C
HAdV-40
Fold Activation of (Gal4)6-Luc
1000
800
600
400
200
0
ns
ns
ns
ns
*
*
Full Length
MYND
No MYND

(Figure 3.17). In cells transfected with HAdV-5 E1A, introducing 0.25µg of full length BS69 resulted in a 44.2% decrease in luciferase activity (Figure 3.17A). Furthermore, cells co-transfected with HAdV-5 E1A and 0.05µg or 0.25µg of MYND showed a 53.5% and 82.2% decrease in luciferase activity respectively compared to cells not transfected with MYND. Cells transfected with HAdV-12 E1A showed higher sensitivity to increasing concentrations of full length BS69 and MYND compared to cells transfected with HAdV-5 E1A (Figure 3.17B). When cells were transfected with 0.05µg and 0.25µg full length BS69, luciferase activity dropped by 34.9% and 57.8% respectively. For cells transfected with 0.05µg and 0.25µg MYND, luciferase activity dropped by 64.1% and 84.4% respectively. Lastly, cells transfected with HAdV-40 E1A are the least sensitive to BS69 mediated repression (Figure 3.17C). Luciferase activity was not significantly different when transfected with increasing concentrations of full length BS69. Increasing concentrations of MYND decreased luciferase activity significantly, as transfecting 0.05µg and 0.25µg of MYND decreased luciferase activity by 34.4% and 63.5% respectively.

E1A from HAdV-12 was most sensitive to BS69 mediated repression of transactivation, followed by HAdV-5, then HAdV-40. It is interesting to note that this pattern mirrors the relative binding affinity of E1A from these three adenovirus types, in that E1A from HAdV-12 binds most strongly to BS69, followed by HAdV-5, and finally HAdV-40 (Figure 3.3A).

3.2.4 Using siRNA to Knock Down Expression of BS69 in A549 and IMR90 cells

BS69 was recently found to be a specific reader of a histone H3 variant, histone H3.3, specifically trimethylated at Lysine residue 36 (H3.3K36me3) (Guo et al., 2014; Wen et al., 2014). BS69 also downregulates expression of target genes by two known mechanisms: suppression of transcriptional elongation and stimulation of intron retention (Guo et al., 2014; Wen et al., 2014). Target genes have been identified by RNA-seq analysis, and verified using qPCR and western blotting (Guo et al., 2014; Wen et al., 2014). To verify the effects of BS69 on gene regulation, we first set out to knock down BS69 expression in

IMR90 and A549 cells. We chose to use IMR90 primary lung fibroblasts and the A549 lung epithelial carcinoma cell line for this study because both cell types are susceptible to adenovirus infection and are widely used for these types of experiments.

We transfected IMR90 and A549 cells using Silencer Select siRNAs (Thermo Fisher siRNA ID: s21153 and s21154) at a final concentration of 10nM 24 hours after being seeded into 6 well plates (Figure 3.18A). siRNA transfections were carried out using SiLenFect Lipid Reagent (BioRad) according to manufacturer's protocol. Cells were then harvested 12, 24, and 48 hours post-transfection and analyzed by western blot. The siRNA knockdown was unsuccessful, as BS69 protein levels did not change with either of the transfected cells regardless of when the cells were harvested. Thus, we repeated the experiment with the following modifications: we increased the concentration of the siRNAs from 10nM to 20nM, combined both siRNAs in one of the samples, transfected A549 cells 4 hours after being seeded, harvested the cells 48 hours post-transfection, and used an siRNA for another target, PKA, as a positive control (Figure 3.18B). The PKA siRNA knockdown was successful, but BS69 protein levels still did not change. We repeated the experiment again using siRNA from another source, Dharmacon (SmartPool siRNA against ZMYND11), which contains four siRNA sequences as opposed to one from Silencer Select. We used 10nM or 20nM final concentration, transfected A549 cells 4 hours after being seeded, and harvested the cells 24 or 48 hours post transfection (Figure 3.18C). BS69 protein levels still did not change using any of these parameters.

To ensure that my inability to detect knockdown of BS69 by siRNA was not related to a problem with antibody specificity, we transfected HT1080 cells with 0.5μg and 2μg of BS69-HA (Figure 3.19A). BS69 protein expression as detected by western blot with this anti-BS69 antibody increased as expected with increasing concentrations of transfected BS69. This suggests that the antibody is indeed specific for BS69. Next, we quantified BS69 mRNA transcript levels after siRNA knockdown using qPCR (Figure 3.19B). We transfected A549 cells with 20nM siRNA 4 hours after being seeded and harvested the cells 48 hours post-transfection. Total RNA was extracted using the PureLink RNA Mini Kit (Thermo Fisher), then converted into cDNA using SuperScript VILO cDNA Synthesis Kit

Figure 3.18. siRNA knockdown of BS69 in IMR90 and A549 cells. A) IMR90 and A549 cells were transfected with 10nM of Silencer Select siRNA (siBS69 A: s21153; siBS69 B: s21154; Thermo Fisher) or a scrambled negative control (Silencer Negative Control No.2, Ambion) 24 hours after being seeded, and harvested 12, 24, and 48 hours post-transfection. Cell lysates were run on a western blot and the membrane was probed with anti-BS69 and anti-actin antibodies. B) A549 cells were transfected with 20nM total of Silencer Select siRNA (siBS69 A: s21153; siBS69 B: s21154; siPKA: s286; Thermo Fisher) or a scrambled negative control (Silencer Negative Control No.2, Ambion) 4 hours after being seeded, and harvested 48 hours post-transfection. Cell lysates were run on a western blot and the membrane was probed with anti-BS69, anti-PKA, and anti-tubulin antibodies. C) A549 cells were transfected with 10nM or 20nM total of Dharmacon siRNA (SmartPool siRNA against ZMYND11) or a scrambled negative control (Silencer Negative Control No.2, Ambion) 4 hours after being seeded, and harvested 24 or 48 hours post-transfection. Cell lysates were run on a western blot and the membrane was probed with anti-BS69, and anti-actin antibodies.

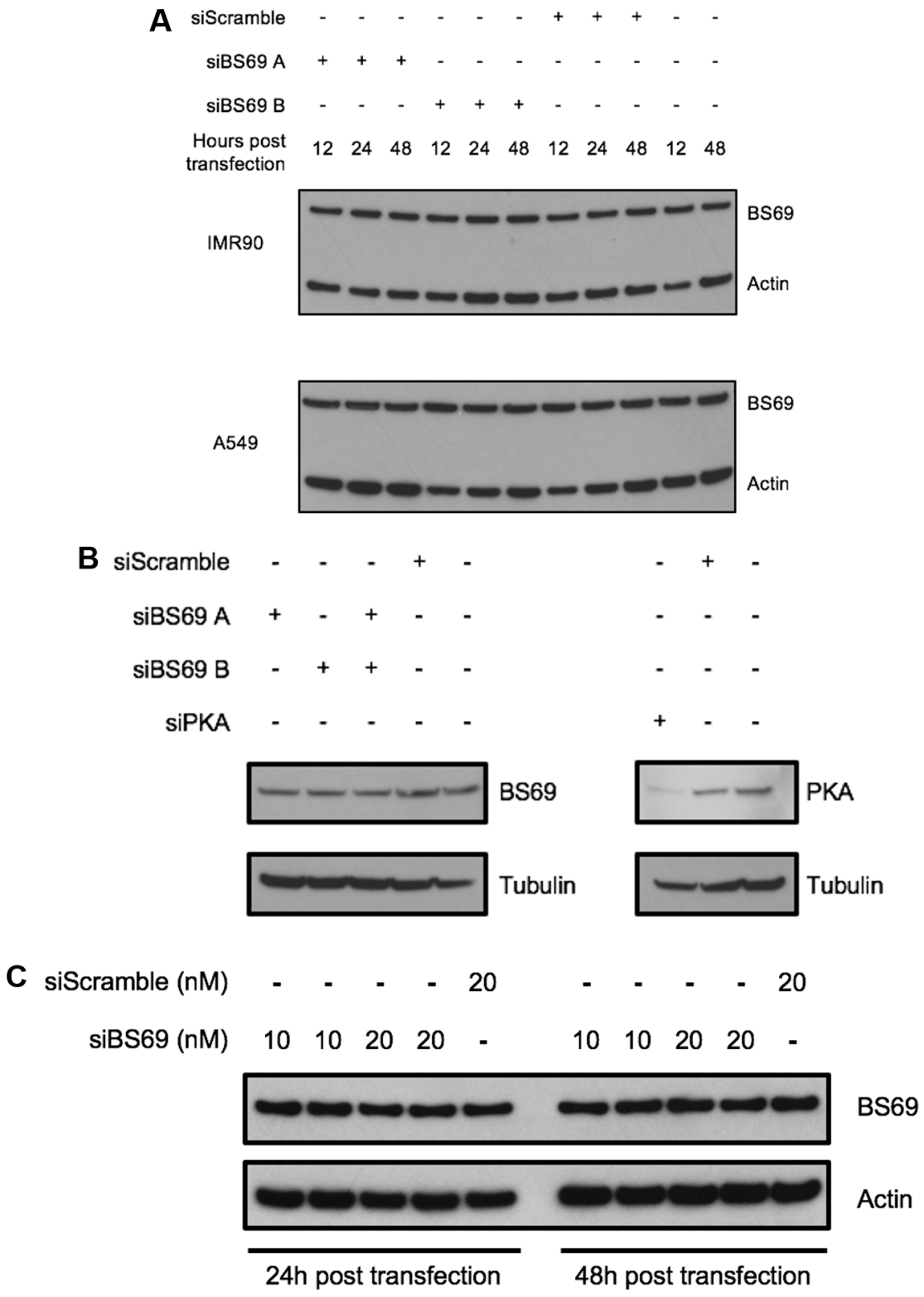
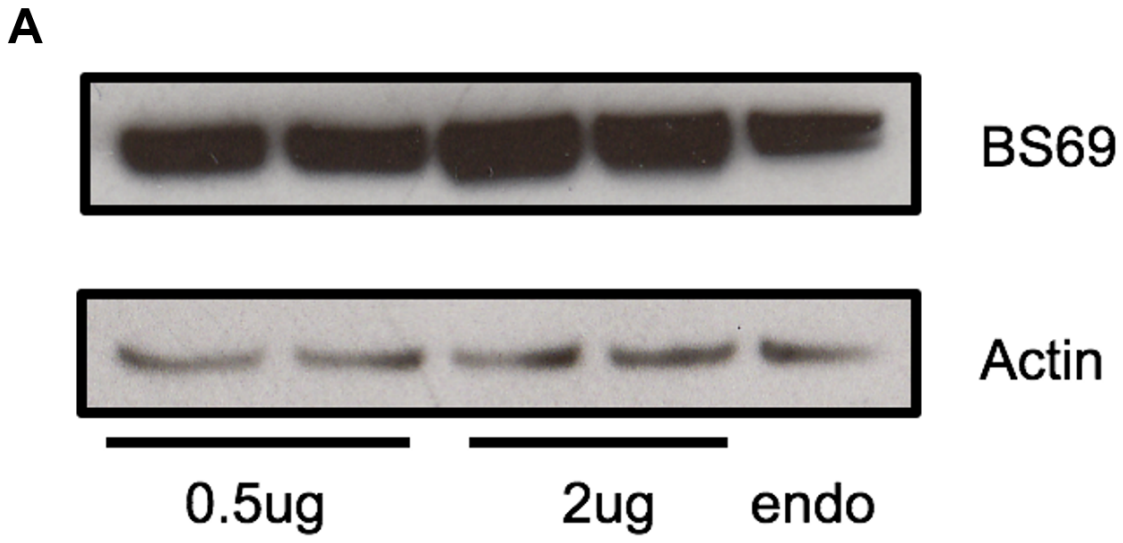
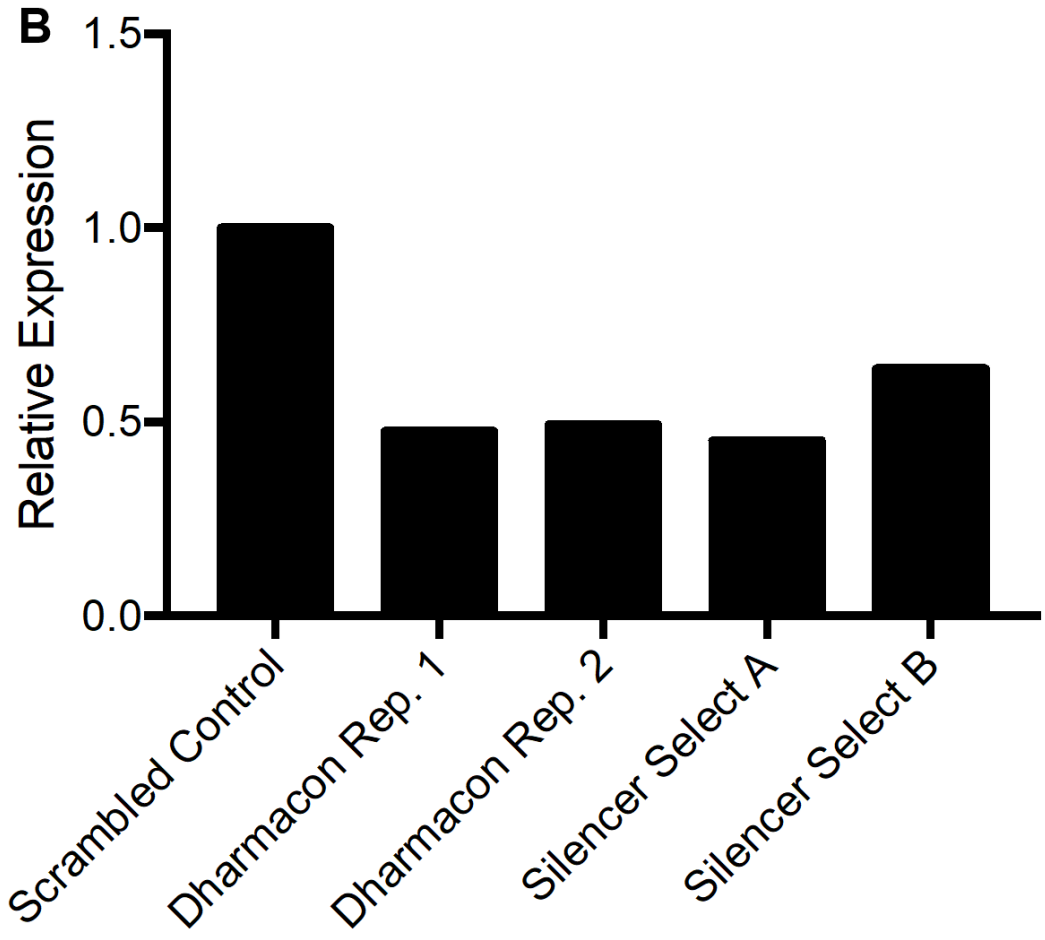

A
siScramble - - - - - - + + + - -
siBS69 A + + + - - - - - - - -
siBS69 B - - - + + + - - - - -
Hours post transfection 12 24 48 12 24 48 12 24 48 12 48
IMR90
BS69
Actin
A549
BS69
Actin
B
siScramble - - - + - - + -
siBS69 A + - + - - - - -
siBS69 B - + + - - - - -
siPKA - - - - - + - -
BS69
PKA
Tubulin
Tubulin
C
siScramble (nM) - - - - 20 - - - - 20
siBS69 (nM) 10 10 20 20 - 10 10 20 20 -
BS69
Actin
24h post transfection
48h post transfection

Figure 3.19. Verification of BS69 antibody specificity and BS69 siRNA induced mRNA transcript knockdown. A) HT1080 cells were transfected with 0.5μg and 2μg of HA-tagged BS69 24 hours after being seeded. Cell lysates were collected 24 hours post transfection. Samples were run on a western blot and the membrane was probed with anti-BS69 and anti-actin antibodies. B) A549 cells were transfected with 20nM total of Silencer Select siRNA (siBS69 A: s21153; siBS69 B: s21154; Thermo Fisher), Dharmacon siRNA (SmartPool siRNA against ZMYND11), or a scrambled negative control (Silencer Negative Control No.2, Ambion) 4 hours after being seeded, and harvested 48 hours post-transfection. Total RNA was extracted and 1μg was used reverse transcribed into cDNA. BS69 transcript levels were quantified using qPCR with SYBR Green. Results were analyzed using the $2^{-\Delta\Delta CT}$ method, using GAPDH as the internal control and normalizing data to that of cells transfected to scrambled siRNA. Results are shown as values from a single experiment using technical triplicates.


A
BS69
Actin
0.5ug
2ug
endo


B
Relative Expression
1.5
1.0
0.5
0.0
Scrambled Control
Dharmacon Rep. 1
Dharmacon Rep. 2
Silencer Select A
Silencer Select B

(Thermo Fisher) according to manufacturer's protocol. cDNA quantification was completed using Power SYBR Green PCR Master Mix kit (Thermo Fisher). Results were obtained using the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001), using GAPDH as the internal control and normalizing the data to that of cells transfected with the scrambled siRNA control. All samples transfected with BS69 siRNA showed an approximate 50% decrease in mRNA transcript levels. Thus, the siRNA knockdown is indeed working to decrease BS69 transcript levels. These results suggest that a long half-life and subsequent low protein turnover rate of BS69 may be the reason why I could not knockdown BS69 protein levels using siRNA.

4 Discussion

4.1 Mapping the Interaction between E1A and BS69

4.1.1 Using Homology Modelling to Construct a Structural Prediction of the MYND Domain of BS69

Homology modelling and protein-ligand docking were used to visualize the molecular determinants of the interaction between E1A and BS69. The model was developed to aid in the design of an E1A mutant incapable of binding with BS69. Within a few months after the construction of our homology model, another group published an NMR structure of the BS69 MYND domain interacting with a peptide from the EBV oncoprotein, EBNA2 (Harter et al., 2016). The EBNA2 peptide contains significant sequence similarity to the PXLXP motif in E1A. Ultimately, both the NMR structure and the homology model were used during the design of our E1A point mutants (Figure 3.7A).

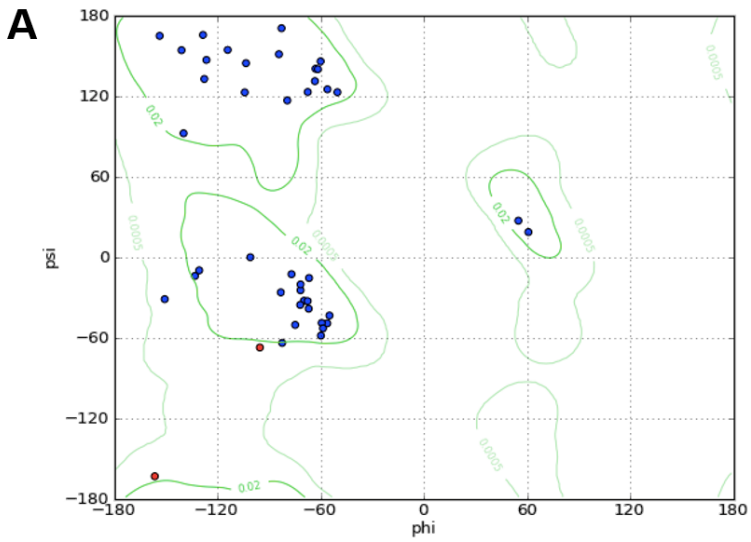
Homology modelling with protein-ligand docking is an alternative method to protein co-crystallization or NMR in determining the binding surface of a protein-peptide interaction. Using this *in silico* method is advantageous in that results may be obtained relatively quickly, so high-throughput modelling is possible. This method may also be used when working with structures that are challenging to produce using either protein crystallization or NMR. Accuracy of these *in silico* models is also consistently improving with every iteration of modelling and docking software (Kozakov et al., 2013). However, any conclusions drawn about the interaction surface from computer modelling must still be verified using other experimental techniques.

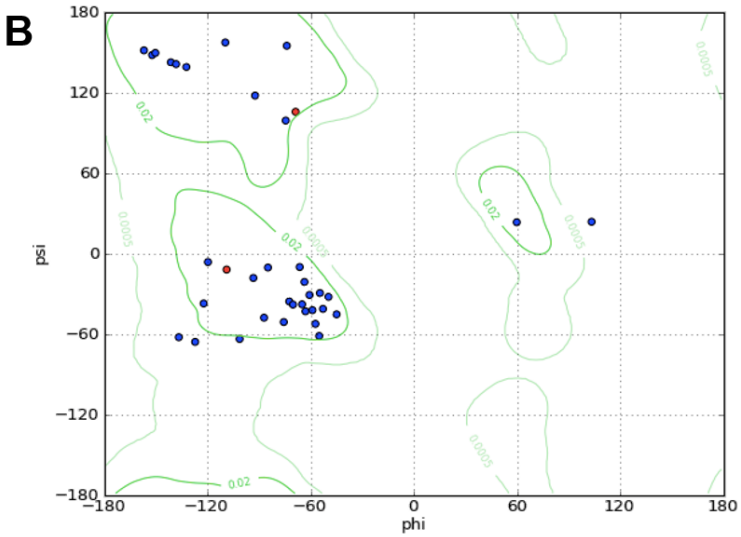
One of the challenges in constructing the BS69 MYND – E1A model was choosing the appropriate template, as every member of the ZMYND family of proteins contain two amino acids between cysteine residues c1b and c2b except for BS69 and ZMYND8, which does not currently have a published structure (Figure 3.1A). These two amino acids between the cysteine residues may function as a spacer, so that the cysteine side chains can more easily achieve the necessary conformation to coordinate a zinc ion and achieve proper

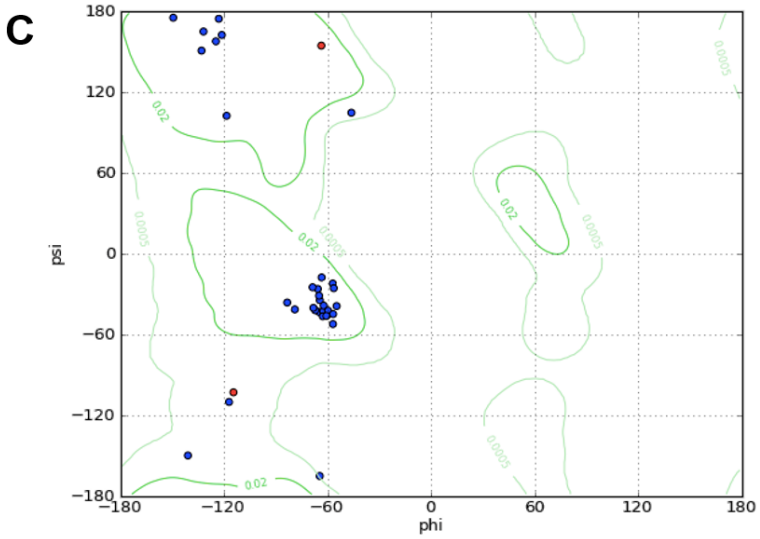
protein folding. This phenomenon is seen in the three MYND domain models shown in Figure 12, where the dihedral angles for the cysteine and surrounding residues are more strained for the BS69 MYND NMR structure and the homology structure compared to ZMYND5, which has the two spacer residues (Figure 4.1). The contour lines on this Ramachandran plot denote the boundary by which 98% or 99.5% of 81,213 non-glycine, non-proline, and non-pre-proline residues from a library of 500 proteins reside (Figure 4.1) (Lovell et al., 2003). Residues within the 98% contour line are "favoured", while residues inside the 99.5% contour line are "allowed"; residues outside of the 99.5% contour line are disallowed due to steric hindrance, unless the residues are glycine, proline, or pre-proline. On the Ramachandran plot, both cysteine residues in the ZMYND5 structure are within the "favoured" region, while only one cysteine residue is within the "favoured" region for the homology model. In contrast, and both cysteine residues are outside the "favoured" region for the BS69 MYND NMR structure. Although all residues within all three structures have dihedral angles that are sterically allowed, the BS69 MYND domain is more strained in this region than ZMYND5 possibly due to the presence of the spacer residues contributing to higher flexibility.

There were two major discrepancies between the computer-generated BS69 MYND model and the NMR model: the leucine sidechain in the PXLXP motif protrudes away from the MYND domain in the homology model, while this sidechain points toward the MYND domain in the NMR model; and the E1A peptide in the homology model is in the opposite orientation compared to the EBNA2 peptide of the NMR model (Figure 3.2). Interestingly, there are a few examples of protein-protein interactions being compatible in opposite orientations such as in Src Src homology 3 (SH3) domain and SUMO-1 binding to their respective binding partners (Feng et al., 1994; Song et al., 2005). Although the PXLXP motif is palindromic, it is unlikely that the E1A-BS69 peptide-protein interaction can occur in both N-C and C-N orientations as this phenomenon is exceedingly rare. The L115 residue is experimentally determined to be important for this interaction, both by myself and others, so it is more likely that the side chain of this residue would be facing towards the MYND domain to facilitate this interaction (Ansieau & Leutz, 2002; Harter et al., 2016). The differences between the *in silico* model and the NMR structures may be caused

Figure 4.1. Ramachandran plots of the MYND domain structures of ZMYND5 and BS69. Ramachandran plots of the A) NMR structure of the BS69 MYND domain (PDB: 5HDA), B) NMR structure of ZMYND5 MYND domain (PDB: 2ODD), and C) homology model of BS69 MYND domain (Harter et al., 2016; Liu et al., 2007). Points in red denote the dihedral angles of the cysteine residues that either have the two spacer residues (NMR model of ZMYND5), or do not have the two spacer residues (NMR or homology model of BS69 MYND). Contour lines in green show the boundary in which residues are “favoured” (within 0.02, or 98%), or “allowed” (within 0.0005, or 99.5%).


A
180
120
60
0
−60
−120
−180
psi
−180
−120
−60
0
60
120
180
phi
0.02
0.0005


B
180
120
60
0
−60
−120
−180
psi
−180
−120
−60
0
60
120
180
phi
0.02
0.0005


C
180
120
60
0
−60
−120
−180
psi
−180
−120
−60
0
60
120
180
phi
0.02
0.0005

by several factors. The Cluspro protein-ligand docking webserver uses rigid-body docking, in that the bond angles and lengths of the receptor, or BS69 MYND, are fixed (Kozakov et al., 2017). This method of docking may not be sufficient for some interactions, as some proteins exhibit slight shifts in conformation upon binding to a ligand. Additionally, the E1A peptide was docked to a homology model of the BS69 MYND domain, which may not have been generated in a conformation conducive to binding with E1A. Flexible docking, where the receptor may undergo conformational changes, may ameliorate some of these problems, but this procedure takes considerably more computing time (Cerqueira et al., 2009).

Although I referred primarily to the BS69-EBNA2 model during the design of our E1A mutants, homology modelling with protein-ligand docking still served as a valuable tool to validate various aspects of the BS69-E1A interaction surface. For example, the leftmost proline residue in all three structures show similar bond angles and spatial orientation, with the side chain protruding towards the MYND domain (Figure 3.2). The consistency between the models is a hint that this proline residue may play an important role in the peptide-protein interaction. The yeast two-hybrid assays using the alanine scanning panel confirm that this residue is crucial in the E1A-BS69 interaction (Figure 3.6). Thus, homology modelling is helpful in both predicting and validating the molecular determinants of a protein-peptide interaction.

4.1.2 Determining the Conservation of Interaction Between BS69 and E1A of Human Adenovirus Species A-F

Using yeast two-hybrid assays, I determined that E1A from HAdV-5, 9, and 12 of species C, D, and A respectively can bind to BS69, while E1A from HAdV-3, 4, and 40 of species B, E, and F cannot (Figure 3.3A). It is unknown if E1A from HAdV-52, the newest characterized adenovirus type from species G, is able to interact with BS69, but it is unlikely as E1A from that species does not contain the PXLXP motif (Figure 1.3). Thus, the interaction between E1A and BS69 is fairly conserved amongst different adenovirus species, and may have an important functional role during an adenovirus infection.

The PXLXP motif was identified by the Leutz lab through sequence comparison and co-IP assays of BS69 with E1A and EBNA2 (Ansieau & Leutz, 2002). As expected, this motif is present in E1A proteins able to bind BS69, and absent or truncated in E1A proteins from HAdV species unable to bind BS69 (Figure 3.3B). Unexpectedly, the sequence alignment shows that the location of the PXLXP motif in HAdV-9 E1A is located upstream compared to the motif of HAdV-5 E1A, yet both proteins can bind BS69 (Figure 3.3B). HAdV-12 E1A also contains two PXLXP motifs (Figure 1.3), but the C-terminal motif alone is sufficient to mediate an interaction with BS69. The HAdV-12 E1A construct I used is missing the first proline residue of the N-terminal motif (Figure 3.3B). EBNA2 from EBV contains two PXLXP motifs in CR7 and CR8 spanning residues 377-391 and residues 437-450 respectively, and both these motifs can bind to BS69 (Harter et al., 2016). It is not known if both PXLXP motifs within HAdV-12 can coordinate interactions with two BS69 molecules concurrently. However, this is unlikely because the motifs are very close together, with only 10 amino acids between the motifs (Figure 1.3). Evidently, the location of the PXLXP motif within the protein is not important in its ability to facilitate an interaction with BS69, and my observations highlight the modularity of the SLiMs within these two viral proteins.

4.1.3 Defining the Minimal Interacting Region and Specific Residues of E1A Required to Bind with BS69

By using the E1A truncation panel in a series of yeast two-hybrid assays, I found that E1A residues 112-119, with the sequence MPNLVPEV, is the minimal interacting region required to bind strongly to BS69. Furthermore, using an alanine scanning mutagenesis panel, I found that E1A residues P113 and L115 are crucial in mediating a strong interaction (Figure 3.6B). I also found that the S111A mutation also resulted in modestly decreased, albeit not significant, interaction strength between E1A and BS69. This observation is also present in the truncation mutation panel, as the peptides containing the N-terminal serine residue exhibit higher binding activity compared to the peptides with the serine truncated (Figure 3.4A). Thus, I have identified three residues within E1A CR2 that are important for interacting with BS69.

To facilitate the truncation mutation analysis, a short peptide linker was used to bridge the GAL4 DBD and the E1A truncation mutants in the bait protein (Figure 3.5). We chose to use SGG as the linker sequence because glycine residues are more flexible compared to other amino acids due to their lack of a β-carbon, while serine was chosen for the small side chain and hydroxyl group to enhance solvation. The linker was included to avoid steric effects between the GAL4 DBD and the BS69 MYND domain that could potentially hinder the interaction between the E1A peptide and BS69 in this system. I have shown that adding the linker increases β-galactosidase activity of our E1A constructs by approximately 100 Miller units (Figure 3.5). The length of the linker is also sufficient, as the first significant drop in activity was observed when a C-terminal residue of the truncation mutant was removed (T4 and T5, Figure 3.4A). Thus, the decreased in activity was not caused by BS69 shifting closer the GAL4 DBD as the E1A peptide was being truncated.

Most of the mutations generated by alanine scanning did not have a significant impact on the binding affinity between E1A and BS69 (Figure 3.6). Yet, truncating some of these residues, such as V119 and M112, caused the binding affinity between E1A and BS69 to significantly decrease (Figure 3.4A). This may be due to these valine and methionine residues playing a role in binding with BS69, but not requiring high sequence stringency. From these observations, alanine scanning mutagenesis is not sufficient in determining all the residues involved in a protein-peptide interaction. Alanine residues can also mediate van der Waals interactions similarly to methionine and valine residues, so mutating these residues to alanine may not cause a detectible change in binding affinity. Although glycine residues are less capable than alanine in facilitating these intermolecular interactions, the lack of a β-carbon enables glycine residues to form conformations that other amino acids cannot; thus, glycine was not used in this mutation panel.

Double and triple E1A point mutants were constructed to improve upon the previously published L115A mutant, as the L115A construct shows some residual interaction with the MYND domain of BS69 in our luciferase assays (Figure 3.16A) (Ansieau & Leutz, 2002). The amino acids being targeted were chosen by referring to the NMR structure of EBNA2 in complex with the MYND domain of BS69 and our E1A-BS69 homology model (Harter et al., 2016). E1A residues M112, P113, L115, and P117 were targeted for mutation, as the

equivalent EBNA2 residues protrude towards the MYND domain in the NMR structure. Because all our double and triple point mutants showed no activity, we ran a western blot to ensure the bait and prey proteins were being sufficiently expressed (Figure 3.7C). We found that all proteins in all samples were expressed as expected, except for the M12 bait protein. Therefore, a more sensitive assay should be used for future studies to measure the interaction between these mutants and BS69. The M13 mutant is a good candidate to proceed with future experimentation, as this construct expresses well and contains the P113A and L115A mutations – both of which were separately identified to decrease the binding affinity between E1A and BS69 using the alanine scanning yeast two-hybrid assays (Figure 3.6, 3.7).

In summary, I found that E1A residues P113 and L115 are strongly involved in the interaction with BS69, while residues S111, M112, and V119 have an ancillary role. Unexpectedly, mutating P117, which is the C-terminal proline of the PXLXP motif, did not change the interaction strength between E1A and BS69 (Figure 3.14B). Although we found several amino acids to target in future studies, it is important to ensure that these residues are not involved in neighbouring E1A interactors. The EVIDLT and LXCXE motifs used by E1A to interact with UBC9 and pRb/STING respectively are downstream of the PXLXP motif (Figure 1.6) and there are no known interaction motifs immediately upstream of the PXLXP motif. These neighbouring interacting partners may also be competing with BS69 to bind E1A CR2 in the truncation mutation assays, resulting in decreased activity of the CR2 control sample compared to E1A truncation mutants that do not contain the EVIDLT and LXCXE motifs (Figure 3.4A). The M13 double point mutant is a promising candidate for future experiments, as there are only two residues targeted for mutation, the mutations are conservative, and the targets are closer to the N-terminus and therefore more distant from the EVIDLT and LXCXE motifs.

4.1.4 Mapping the E1A-BS69 Binding Surface Using a Peptide Array

Peptide arrays were used to further characterize the interaction surface between BS69 and E1A. By using positional scanning, we mutated every position on the E1A peptide to every other residue. This technique allows us to precisely probe the BS69 minimal interacting region to determine the sequence stringency needed to facilitate a strong interaction with BS69. Unlike yeast two-hybrid assays, peptide arrays use purified proteins and peptides, and can thus measure direct protein-peptide interactions. Peptide arrays are also an especially powerful tool for studying E1A interactions, as E1A is intrinsically disordered in all but CR3 and is composed of densely-packed interaction motifs (Figure 1.6). Although these arrays allow for high-throughput screening of a protein-peptide interaction, there are also several constraints. Yeast two-hybrid and co-IP assays can both measure interactions between two globular proteins, while peptide arrays are restricted to protein-peptide interactions. Peptide arrays are also limited in that the peptides may only be up to a maximum length, approximately 15 residues, before sequence fidelity declines due to the compounding effect of peptide synthesis efficiencies. The protein-peptide interaction also takes place in PBS, which does not accurately mimic the constituents of a cellular environment. The high concentration of the proteins and peptides used in a peptide array experiment are also a concern, as signals detected from a peptide array may not accurately depict the interactions formed in a cellular environment. Thus, results of a peptide array must be verified using other more biologically relevant techniques, such as co-IP. Nevertheless, peptide arrays are a powerful and flexible tool to comprehensively and systematically study protein-peptide interactions, and are especially suitable for intrinsically disordered proteins containing SLiMs, such as E1A.

Protein purification of GST- and 6xHis- tagged BS69 were completed using affinity chromatography. I have also purified GST-tagged pRb and UBC9, which are E1A binding partners that interact with SLiMs neighbouring the PXLXP motif, in preparation for future peptide array experiments (Figure 3.10). The BS69 MYND domain was chosen to be our target to purify, as my preliminary experiments showed that purifying the MYND domain instead of full length BS69 resulted in fewer degradation products (not shown). Coomassie

Blue staining of purified 6xHis tagged BS69 MYND revealed protein bands at 75kDa, 50kDa, 25 kDa, and 20 kDa, with the 25kDa band being the most prominent (Figure 3.8A). The 75kDa and 50kDa bands reflect polymers of the MYND domain, while the 20kDa band could be a degradation product. A western blot of the same sample showed a similar laddering effect, with bands at 25kDa, 50 kDa, 75 kDa, 100kDa, and 125 kDa, which further affirms the presence of protein polymers or aggregates (Figure 3.8B). Aggregation and degradation may have been caused by a shortage of zinc in the media, as each molecule of BS69 MYND coordinates two zinc ions. Lack of zinc may have caused multiple molecules of MYND to share zinc ions, or caused degradation due to protein misfolding. To ameliorate this problem, I supplemented the growth media with 100μM ZnCl during the next round of protein purification. Addition of ZnCl did not alter the purification profile of 6xHis-tagged BS69, and may have negatively impacted purification by competing with the Nickel-NTA substrate to bind the 6xHis tag (Figure 3.9). I concurrently purified GST-tagged BS69 MYND, which appeared as a single band at 47kDa in the Coomassie Blue gel (Figure 3.9). Due to time constraints, 6xHis-tagged BS69 was used to probe the peptide array, but future experiments should be conducted using GST-tagged BS69 MYND due to better protein purity, lack of protein aggregation, and fewer degradation products.

Background signal was very high for the peptide array experiments using peptides both HAdV-5 and HAdV-12 E1A. E1A from HAdV-12 can also bind BS69, so we chose to use this E1A both as a positive control and to determine if the interaction mechanism is similar in other adenovirus species. We expected the wildtype peptide sequences (circled in red) to produce the strongest signal, but these spots displayed similar intensities to most other spots on the membrane (Figure 3.11). Unexpectedly, peptides with residues sequentially mutated to arginine produced the strongest signal in both HAdV-5 and HAdV-12 E1A arrays when probed with anti-His and anti-BS69 antibodies (Figure 3.11, 3.12). These signals were not present when the membranes were stripped and re-probed with anti-His antibody without first incubating the array with purified His-MYND (Figure 3.13). Thus, these signals may have resulted from nonspecific binding of the protein to the peptides on the membrane. The anti-His antibody also contributed to some background signal, shown by the image of the stripped membrane (Figure 3.13). In future experiments, the GST-tagged BS69 MYND should be used in place of His-tagged BS69 to improve the quality

of purified protein, and more stringent wash protocols should also be used to minimalize background signals. The sequence fidelity of the peptides synthesized on the membrane should also be evaluated.

4.1.5 Verifying the E1A-BS69 Interaction Using Co-Immunoprecipitation Assays

I verified the interaction between E1A and BS69 in a mammalian cell culture system using co-IP assays. Although E1A CR2 was sufficient to form an interaction between BS69, full length E1A pulls down BS69 more efficiently (Figure 3.14). Thus, there appears to be other regions within E1A used to facilitate the interaction with BS69. There is conflicting data in the literature about the involvement of CR3 in this interaction. The Bernards group found that CR3 is the main region by which E1A uses to facilitate its interaction with BS69 (Hateboer et al., 1995). However, the Leutz group found that 13S and 12S E1A both bound BS69 equally as well (Ansieau & Leutz, 2002). The Bernards group used murine BS69, while the Leutz group used human BS69 in their respective co-IP experiments. BS69 is highly conserved between humans and mice, with 14 amino acid changes in total between these species (Figure 1.9). The BS69 MYND domain sequence is identical in human and mice, and most of the variation is located between the PWWP domain the and the MYND domain. It is possible that regions other than the MYND domain of murine BS69 interacts with CR3 of E1A, but the same regions in human BS69 does not interact with E1A. To determine if CR3 enhances the ability of E1A to bind to BS69, I created the CR2-CR3 E1A truncation mutant. Unexpectedly, this E1A construct showed decreased binding strength compared to both CR2 and full length E1A (Figure 3.14) This observation may be caused by improper protein expression of this E1A construct, or misfolding of CR3.

Co-IP assays should also be used to verify the interaction data obtained from the previous yeast two-hybrid experiments. Yeast two-hybrid assays are more sensitive than co-IP in that two-hybrid assays can detect transient interactions, and the proteins of interest in two-hybrid assays are relocalized to the nucleus. Thus, two-hybrid assays may detect interactions that would not normally occur under mammalian physiological conditions due

to differences in subcellular localization or binding affinity. Optimally, co-IP assays should be conducted using endogenous proteins instead of using a vector to express the proteins of interest. The proteins I expressed in the co-IP assays are tagged with either HA or GFP, and these tags may alter the binding affinity between the E1A constructs and BS69. Furthermore, overexpressing proteins from a plasmid alters the concentrations and stoichiometric balance of the proteins of interest. Thus, co-IP assays using overexpressed proteins may also produce interactions that do not occur under physiological conditions. In future experiments, it would be prudent to use co-IP assays with endogenous proteins, with E1A mutants introduced via adenovirus infections. Mutant adenovirus may also display altered growth and infectivity, so precautions must be taken to control these variables. Select single and double E1A point mutants showing altered binding affinity to BS69, such as M1, M3, M5, and M13, should be included in these future interaction studies to verify these observations in a mammalian cell culture system.

4.2 Analysis of BS69 Mediated Changes in E1A Transactivation

4.2.1 Determining the Region of BS69 Responsible for Changes in E1A Mediated Transactivation

Using luciferase assays in HT1080 cells, I have shown that BS69 decreases E1A mediated transactivation in a dose dependent manner. These observations agree with the results published by the Bernards group (Figure 3.15) (Hateboer et al., 1995). I have also shown that this repression of E1A transactivation requires an intact PXLXP motif, as transactivation by the E1A L115A mutant was not significantly changed by increasing concentrations of BS69 (Figure 3.15). This observation is consistent with the results found by the Leutz lab, although their experiments were performed in the opposite orientation whereby E1A was used to neutralize BS69-mediated gene repression (Ansieau & Leutz, 2002). With increasing concentrations of BS69, the luciferase activity of the L115A mutant also steadily decreased, albeit not significantly. Thus, there may be another region in E1A that interacts with BS69, or the L115A mutant does not completely abrogate binding

between CR2 and the MYND domain of BS69. The use of the double mutant, such as P113A and L115A, that exhibits a more complete binding defect in this assay may help resolve this question.

To determine if the MYND domain of BS69 was necessary and sufficient to repress E1A transactivation, I repeated the previous experiments and added increasing concentrations of BS69 MYND (residues 427-602) and BS69 No MYND (residues 46-426) constructs (Figure 3.16). Transfecting cells with 0.95μg of BS69, regardless of construct, resulted in significantly decreased luciferase activity in almost all samples. This phenomenon is likely due to a technical error during the experiment causing protein overexpression, so discussion will proceed omitting these results. In this set of experiments, L115A E1A displayed overall greater luciferase activity compared to wildtype E1A. This may be caused by endogenous BS69 inhibiting wildtype E1A transactivation, but not influencing transactivation by L115A E1A. As expected, adding increasing concentrations of full length BS69 mirrored the results from our previous experiment, and adding increasing concentrations of No MYND BS69 did not significantly change transactivation ability of both wildtype and L115A E1A. Therefore, the MYND domain of BS69 is essential in BS69-mediated inhibition of E1A transactivation. In samples with wildtype E1A, increasing concentrations of BS69 MYND caused a decrease in E1A mediated transactivation at a faster rate than full length BS69. This is likely due to the higher level of protein expression by the BS69 MYND compared to full length BS69 at the same transfection concentration (Figure 3.16B). The smaller MYND domain by itself may also be more accessible to E1A compared to full length BS69, as the other globular domains of BS69 may either introduce steric hindrance to the promoter region of the reporter plasmid, or be used to bind to other host factors and be recruited away from E1A. Unexpectedly, adding BS69 MYND also caused a decrease in transactivation by the L115A E1A mutant. Again, this observation is likely caused by increased BS69 MYND expression or accessibility to E1A. The L115A mutant binds to BS69 with much less affinity compared to wildtype E1A; however, an interaction may still occur with sufficiently high protein concentrations. To better understand these observations, protein expressions levels should be consistent between the different BS69 constructs in future experiments, and another E1A mutant should be used to completely abrogate the interaction with BS69.

4.2.2 Establishing the Conservation of BS69 Mediated Repression of E1A Transactivation in HAdV-5, 12, and 40

Using luciferase assays and various adenovirus types, I showed that E1A proteins from species with greater binding affinity to BS69 are more susceptible to BS69 mediated repression of E1A transactivation. E1A from HAdV-12 binds with the greatest affinity to BS69, followed by HAdV-5, then HAdV-40 (Figure 3.3). BS69-mediated repression of E1A transactivation follows the same pattern, where E1A from HAdV-12 is most sensitive, followed by HAdV-5 and finally HAdV-40 (Figure 3.17). This phenomenon is most clearly illustrated in samples transfected with full length BS69, where HAdV-12 E1A transactivation is significantly inhibited by low and high concentrations of full length BS69, HAdV-5 E1A is only significantly inhibited by high concentrations of full length BS69, and HAdV-40 E1A is not significantly affected by full length BS69. As expected, increasing levels of BS69 No MYND had no significant effect on transactivation by E1A from any of the tested adenovirus types.

E1A from these adenovirus types also display dramatically different levels of transactivation. When these samples were not co-transfected with BS69, E1A from HAdV-12 produced the highest level of transactivation, followed by HAdV-40, and finally HAdV-5. These observations agree with previously published results from the Mymryk lab (Ablack et al., 2010). E1A CR3 is a zinc finger domain, and can facilitate transactivation on its own when fused to a DBD (Ablack et al., 2010). E1A CR3 is the most conserved domain between different adenovirus species, yet transactivation ability between adenovirus species varies dramatically. This may be caused by different mechanisms between species by which CR3 facilitates transactivation, but this may also be caused by unique ways in which E1A from different adenovirus species are regulated. Here, I have also shown that BS69 differentially affects E1A transactivation from different adenovirus species in a manner that is dependent on the interaction strength between E1A and BS69. Thus, the ability of BS69 to repress E1A dependent transactivation may contribute to some of the observed differences in transactivation between HAdV species.

4.2.3 Knockdown of BS69 in A549 and IMR90 Cells

To proceed with future experiments to study other physiological processes influenced by the E1A-BS69 interaction, such as how E1A affects BS69-mediated gene regulation and how BS69 affects adenovirus growth, it is first necessary to create a cell line that does not express BS69. Human BS69 is ubiquitously expressed in all tissues and localized to the nucleus, and there is no commercially available cell line to our knowledge that does not express BS69 (Velasco et al., 2006). We chose to knock down BS69 in IMR90 primary cells and the A549 cell line due to their susceptibility to adenovirus infection for future experiments. I was unable to knock down BS69 using siRNA from ThermoFisher or Dharmacon (Figure 3.18) in either of the two cell lines. I showed the experimental technique and reagents were working as intended by successfully knocking down PKA (Figure 3.18B), ensuring the antibody was specific to BS69 (Figure 3.19A), and verifying the reduction of BS69 transcript levels (Figure 3.19B). Other groups have attempted to knock down BS69 as well, but most have used short hairpin RNA (shRNA) delivered using a lentiviral vector (Guo et al., 2014; Wen et al., 2014). Another group used siRNA to knock down BS69, but only transcript levels were shown (Mackmull et al., 2015). siRNA knockdowns are short-term, while shRNA knockdown via a lentiviral gene delivery system is long-term. Thus, the reason why my siRNA knockdowns were unsuccessful could be due to BS69 having a long half-life. In future attempts, clustered regularly interspaced short palindromic repeats (CRISPR) genome editing or shRNA should be used to create BS69 knockout cell lines.

4.3 Summary of Findings and Future Directions

This project has revealed several interesting observations about the interaction between E1A and BS69. Firstly, I found that this interaction is conserved in HAdV-5, 9, and 12 of species C, D, and A respectively, and is absent in HAdV-3, 4, and 40 of species B, E, and F respectively. This finding also confirms the importance of the previously described PXLXP motif, as adenoviruses that have this motif within E1A can bind BS69, while adenoviruses that lack the motif are not able to bind BS69 (Ansieau & Leutz, 2002). The

minimal interacting region was also defined as residues 112-119 in HAdV-5 E1A, with P113 and L115 having the most important roles in this interaction. S111, M112, and V119 of HAdV-5 E1A also contribute to this interaction, but these residues play an ancillary role. Additionally, I found that full length E1A binds to BS69 with greater affinity than CR2 alone. However, it is still unclear if CR3 plays a role in the interaction between E1A and BS69 despite conflicting data in the literature (Ansieau & Leutz, 2002; Hateboer et al., 1995). A future direction of this interaction study would be to probe the remaining regions of E1A to discover other BS69 interacting sites. Furthermore, more experiments should be completed to determine the relative binding affinity of the E1A double and triple point mutants with BS69, as well as neighbouring binding partners such as pRb, UBC9, and STING. Peptide arrays are a powerful tool that may accelerate this process, with the benefit of confirming that the interaction between E1A and BS69 is direct. These experiments will ultimately help find an E1A mutant that completely and cleanly abrogates binding with BS69 without perturbing the interaction with other target proteins.

Consistent with previous results, I have also confirmed that BS69 is a potent inhibitor of E1A transactivation, and that this function of BS69 is dependent on the fidelity of the PXLXP motif on E1A (Ansieau & Leutz, 2002). Inhibition by BS69 is also dose-dependent, as higher concentrations of BS69 resulted in lower transactivation ability of E1A. The C-terminal MYND domain of BS69 is both necessary and sufficient to inhibit E1A transactivation. I have also found that BS69-mediated inhibition of E1A transactivation is conserved in HAdV-5, 12, and, 40, with HAdV-5 having the greatest sensitivity and HAdV-40 being the least sensitive. It was surprising to find that transactivation by HAdV-40 E1A was affected by BS69, as CR2 from this E1A did not bind to BS69 according to my yeast two-hybrid experiments. Thus, other regions of E1A may facilitate the interaction between E1A and BS69 in HAdV-40, or the high concentration of the BS69 MYND construct within this luciferase system resulted in an experimental artefact. In future experiments, it would be prudent to use a double or triple E1A point mutant in these luciferase assays, and to ensure that the expression levels of the different BS69 truncation mutants are more consistent.

With the successful mapping of an interaction site between E1A and BS69, a future direction of this project involves the generation of a cell line that does not express BS69. Using siRNA to knock down BS69, which has a low protein turnover rate, is not effective, so other methods should be used instead. For example, using CRISPR or shRNA may yield more fruitful results, as these methods are permanent while the effects of siRNA are transient. After generating a BS69 knockout cell line, we will be able to measure the impact of the E1A-BS69 interaction on viral growth and regulation of host and viral gene expression. BS69 is a histone reader that downregulates genes that are decorated by histone H3.3K36me3 (Guo et al., 2014; Wen et al., 2014). Other studies have also shown that BS69 fused to GAL4 DBD represses transcription in a luciferase system, and that E1A neutralizes BS69-mediated gene repression (Masselink & Bernards, 2000). It would be interesting to see if E1A is able to abrogate BS69-mediated gene repression at the epigenetic level, and to discover the mechanism by which this occurs.

5 References

Ablack, J. N. G., Pelka, P., Yousef, A. F., Turnell, A. S., Grand, R. J. A., & Mymryk, J. S. (2010). Comparison of E1A CR3-dependent transcriptional activation across six different human adenovirus subgroups. *Journal of Virology*, *84*(24), 12771–81.

Adams A, Gottschling DE, Kaiser CA, S. T. (1997). *Methods in Yeast Genetics*. New York: Cold Spring Harbor Press.

Ahi, Y. S., & Mittal, S. K. (2016). Components of Adenovirus Genome Packaging. *Frontiers in Microbiology*, *7*, 1503.

Ahi, Y. S., Vemula, S. V, & Mittal, S. K. (2013). Adenoviral E2 IVa2 protein interacts with L4 33K protein and E2 DNA-binding protein. *The Journal of General Virology*, *94*(Pt 6), 1325–34.

Ahn, S. H., Kim, M., & Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Molecular Cell*, *13*(1), 67–76.

Alevizopoulos, K., Catarin, B., Vlach, J., & Amati, B. (1998). A novel function of adenovirus E1A is required to overcome growth arrest by the CDK2 inhibitor p27(Kip1). *The EMBO Journal*, *17*(20), 5987–97.

Alevizopoulos, K., Sanchez, B., & Amati, B. (2000). Conserved region 2 of adenovirus E1A has a function distinct from pRb binding required to prevent cell cycle arrest by p16INK4a or p27Kip1. *Oncogene*, *19*(16), 2067–74.

Ansieau, S., & Leutz, A. (2002). The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif. *The Journal of Biological Chemistry*, *277*(7), 4906–10.

Atkinson, R. L. (2007). Viruses as an etiology of obesity. *Mayo Clinic Proceedings*,

82(10), 1192–8.

Atkinson, R. L., Dhurandhar, N. V, Allison, D. B., Bowen, R. L., Israel, B. A., Albu, J. B., & Augustus, A. S. (2005). Human adenovirus-36 is associated with increased body weight and paradoxical reduction of serum lipids. *International Journal of Obesity*, *29*(3), 281–6.

Avvakumov, N., Kajon, A. E., Hoeben, R. C., & Mymryk, J. S. (2004). Comprehensive sequence analysis of the E1A proteins of human and simian adenoviruses. *Virology*, *329*(2), 477–92.

Avvakumov, N., Wheeler, R., D'Halluin, J. C., & Mymryk, J. S. (2002). Comparative sequence analysis of the largest E1A proteins of human and simian adenoviruses. *Journal of Virology*, *76*(16), 7968–75. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=155151&tool=pmcentrez&rendertype=abstract

Bagchi, S., Raychaudhuri, P., & Nevins, J. R. (1990). Adenovirus E1A proteins can dissociate heteromeric complexes involving the E2F transcription factor: a novel mechanism for E1A trans-activation. *Cell*, *62*(4), 659–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2143697

Behjati, S., Tarpey, P. S., Presneau, N., Scheipl, S., Pillay, N., Van Loo, P., … Flanagan, A. M. (2013). Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nature Genetics*, *45*(12), 1479–82.

Benedict, C. A., Norris, P. S., Prigozy, T. I., Bodmer, J. L., Mahr, J. A., Garnett, C. T., … Ware, C. F. (2001). Three adenovirus E3 proteins cooperate to evade apoptosis by tumor necrosis factor-related apoptosis-inducing ligand receptor-1 and -2. *The Journal of Biological Chemistry*, *276*(5), 3270–8.

Bennett, E. M., Bennink, J. R., Yewdell, J. W., & Brodsky, F. M. (1999). Cutting edge: adenovirus E19 has two mechanisms for affecting class I MHC expression. *Journal of Immunology*, *162*(9), 5049–52. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/10227971

Berk, A. J. (1986). Adenovirus promoters and E1A transactivation. *Annual Review of Genetics*, *20*(1), 45–79.

Berk, A. J. (2007). Adenoviridae: The Viruses and Their Replication. In D. M. Knipe & P. M. Howley (Eds.), *Fields Vriology* (5th ed., pp. 2355–2394). Lippincott Williams & Wilkins.

Berk, A. J., Lee, F., Harrison, T., Williams, J., & Sharp, P. A. (1979). Pre-early adenovirus 5 gene product regulates synthesis of early viral messenger RNAs. *Cell*, *17*(4), 935–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/487437

Bewley, M. C., Springer, K., Zhang, Y. B., Freimuth, P., & Flanagan, J. M. (1999). Structural analysis of the mechanism of adenovirus binding to its human cellular receptor, CAR. *Science*, *286*(5444), 1579–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10567268

Bortoluzzi, A., Amato, A., Lucas, X., Blank, M., & Ciulli, A. (2017). Structural Basis of Molecular Recognition of Helical Histone H3 Tail by PHD Finger Domains. *The Biochemical Journal*, BCJ20161053.

Boyer, T. G., Martin, M. E., Lees, E., Ricciardi, R. P., & Berk, A. J. (1999). Mammalian Srb/Mediator complex is targeted by adenovirus E1A protein. *Nature*, *399*(6733), 276–9.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*, 248–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/942051

Brestovitsky, A., Nebenzahl-Sharon, K., Kechker, P., Sharf, R., & Kleinberger, T. (2016). The Adenovirus E4orf4 Protein Provides a Novel Mechanism for Inhibition of the DNA Damage Response. *PLoS Pathogens*, *12*(2), e1005420.

Bridge, E., & Ketner, G. (1990). Interaction of adenoviral E4 and E1b products in late gene expression. *Virology*, *174*(2), 345–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2137659

Burgert, H. G., Maryanski, J. L., & Kvist, S. (1987). E3/19K protein of adenovirus type 2 inhibits lysis of cytolytic T lymphocytes by blocking cell-surface expression of histocompatibility class I antigens. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(5), 1356–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2950523

Cantin, G. T., Stevens, J. L., & Berk, A. J. (2003). Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(21), 12003–8.

Carson, C. T., Orazio, N. I., Lee, D. V, Suh, J., Bekker-Jensen, S., Araujo, F. D., … Weitzman, M. D. (2009). Mislocalization of the MRN complex prevents ATR signaling during adenovirus infection. *The EMBO Journal*, *28*(6), 652–62.

Cathomen, T., & Weitzman, M. D. (2000). A functional complex of adenovirus proteins E1B-55kDa and E4orf6 is necessary to modulate the expression level of p53 but not its transcriptional activity. *Journal of Virology*, *74*(23), 11407–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11070042

Cauet, G., Strub, J.-M., Leize, E., Wagner, E., Van Dorsselaer, A., & Lusky, M. (2005). Identification of the glycosylation site of the adenovirus type 5 fiber protein. *Biochemistry*, *44*(14), 5453–60.

Cepko, C. L., & Sharp, P. A. (1982). Assembly of adenovirus major capsid protein is mediated by a nonvirion protein. *Cell*, *31*(2 Pt 1), 407–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7159928

Cerqueira, N. M. F. S. A., Bras, N. F., Fernandes, P. A., & Ramos, M. J. (2009). MADAMM: a multistaged docking with an automated molecular modeling protocol.

Proteins, *74*(1), 192–206.

Challberg, M. D., Desiderio, S. V, & Kelly, T. J. (1980). Adenovirus DNA replication in vitro: characterization of a protein covalently linked to nascent DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(9), 5105–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6933548

Chan, K.-M., Fang, D., Gan, H., Hashizume, R., Yu, C., Schroeder, M., … Zhang, Z. (2013). The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes & Development*, *27*(9), 985–90.

Chang, S.-Y. S.-F., Lee, C.-N., Lin, P.-H., Huang, H.-H., Chang, L.-Y., Ko, W., … Kao, C.-L. (2008). A community-derived outbreak of adenovirus type 3 in children in Taiwan between 2004 and 2005. *Journal of Medical Virology*, *80*(1), 102–12.

Chattopadhyay, D., Ghosh, M. K., Mal, A., & Harter, M. L. (2001). Inactivation of p21 by E1A leads to the induction of apoptosis in DNA-damaged cells. *Journal of Virology*, *75*(20), 9844–56.

Chou, W.-L., Chung, Y.-L., Fang, J.-C., & Lu, C.-A. (2017). Novel interaction between CCR4 and CAF1 in rice CCR4-NOT deadenylase complex. *Plant Molecular Biology*, *93*(1–2), 79–96.

Christensen, J. B., Byrd, S. A., Walker, A. K., Strahler, J. R., Andrews, P. C., & Imperiale, M. J. (2008). Presence of the adenovirus IVa2 protein at a single vertex of the mature virion. *Journal of Virology*, *82*(18), 9086–93.

Chung, C. T., Niemela, S. L., & Miller, R. H. (1989). One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(7), 2172–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2648393

Cobben, J. M., Weiss, M. M., van Dijk, F. S., De Reuver, R., de Kruiff, C., Pondaag, W., … Yntema, H. G. (2014). A de novo mutation in ZMYND11, a candidate gene for 10p15.3 deletion syndrome, is associated with syndromic intellectual disability.

European Journal of Medical Genetics, *57*(11–12), 636–8.

Cobrinik, D. (2005). Pocket proteins and cell cycle control. *Oncogene*, *24*(17), 2796–809.

Coe, B. P., Witherspoon, K., Rosenfeld, J. A., van Bon, B. W. M., Vulto-van Silfhout, A. T., Bosco, P., … Eichler, E. E. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nature Genetics*, *46*(10), 1063–71.

Davey, N. E., Travé, G., & Gibson, T. J. (2011). How viruses hijack cell regulation. *Trends in Biochemical Sciences*, *36*(3), 159–69.

DeCaprio, J. A. (2009). How the Rb tumor suppressor structure and function was revealed by the study of Adenovirus and SV40. *Virology*, *384*(2), 274–84.

DeScipio, C., Conlin, L., Rosenfeld, J., Tepperberg, J., Pasion, R., Patel, A., … Krantz, I. D. (2012). Subtelomeric deletion of chromosome 10p15.3: clinical findings and molecular cytogenetic characterization. *American Journal of Medical Genetics*, *158A*(9), 2152–61.

Dhurandhar, N. V, Israel, B. A., Kolesar, J. M., Mayhew, G. F., Cook, M. E., & Atkinson, R. L. (2000). Increased adiposity in animals due to a human virus. *International Journal of Obesity and Related Metabolic Disorders*, *24*(8), 989–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10951537

Dix, I., & Leppard, K. N. (1995). Expression of adenovirus type 5 E4 Orf2 protein during lytic infection. *The Journal of General Virology*, *76 (Pt 4)*(4), 1051–5.

Dobner, T., Horikoshi, N., Rubenwolf, S., & Shenk, T. (1996). Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. *Science*, *272*(5267), 1470–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8633237

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes & Development*, *12*(15), 2245–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9694791

Eckner, R., Ewen, M. E., Newsome, D., Gerdes, M., DeCaprio, J. A., Lawrence, J. B., & Livingston, D. M. (1994). Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes & Development*, *8*(8), 869–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7523245

Elsing, A., & Burgert, H. G. (1998). The adenovirus E3/10.4K-14.5K proteins down-modulate the apoptosis receptor Fas/Apo-1 by inducing its internalization. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(17), 10072–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9707602

Enders, J. F., Bell, J. A., Dingle, J. H., Francis, T., Hilleman, M. R., Huebner, R. J., & Payne, A. M. (1956). Adenoviruses: group name proposed for new respiratory-tract viruses. *Science*, *124*(3212), 119–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13337359

Evans, J. D., & Hearing, P. (2003). Distinct roles of the Adenovirus E4 ORF3 protein in viral DNA replication and inhibition of genome concatenation. *Journal of Virology*, *77*(9), 5295–304. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12692231

Feng, S., Chen, J. K., Yu, H., Simon, J. A., & Schreiber, S. L. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science*, *266*(5188), 1241–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7526465

Ferguson, B., Krippl, B., Andrisani, O., Jones, N., Westphal, H., & Rosenberg, M. (1985). E1A 13S and 12S mRNA products made in Escherichia coli both function as nucleus-localized transcription activators but do not directly bind DNA. *Molecular and Cellular Biology*, *5*(10), 2653–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2942760

Fessler, S. P., & Young, C. S. (1998). Control of adenovirus early gene expression during the late phase of infection. *Journal of Virology*, *72*(5), 4049–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9557693

Field, J., Gronostajski, R. M., & Hurwitz, J. (1984). Properties of the adenovirus DNA polymerase. *The Journal of Biological Chemistry*, *259*(15), 9487–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6540263

Fontebasso, A. M., Papillon-Cavanagh, S., Schwartzentruber, J., Nikbakht, H., Gerges, N., Fiset, P.-O., … Kieran, M. W. (2014). Recurrent somatic mutations in ACVR1 in pediatric midline high-grade astrocytoma. *Nature Genetics*, *46*(5), 462–6.

Frank, R. (2002). The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *Journal of Immunological Methods*, *267*(1), 13–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12135797

Frese, K. K., Lee, S. S., Thomas, D. L., Latorre, I. J., Weiss, R. S., Glaunsinger, B. A., & Javier, R. T. (2003). Selective PDZ protein-dependent stimulation of phosphatidylinositol 3-kinase by the adenovirus E4-ORF1 oncoprotein. *Oncogene*, *22*(5), 710–21.

Friedman, J. M., & Horwitz, M. S. (2002). Inhibition of tumor necrosis factor alpha-induced NF-kappa B activation by the adenovirus E3-10.4/14.5K complex. *Journal of Virology*, *76*(11), 5515–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11991979

Frolov, M. V, & Dyson, N. J. (2004). Molecular mechanisms of E2F-dependent activation and pRB-mediated repression. *Journal of Cell Science*, *117*(Pt 11), 2173–81.

Fuchs, M., Gerber, J., Drapkin, R., Sif, S., Ikura, T., Ogryzko, V., … Livingston, D. M. (2001). The p400 complex is an essential E1A transformation target. *Cell*, *106*(3), 297–307. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11509179

Fujisawa, T., & Filippakopoulos, P. (2017). Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nature Reviews. Molecular Cell Biology*, *18*(4), 246–262.

Gallimore, P. H. (1972). Tumour production in immunosuppressed rats with cells

transformed in vitro by adenovirus type 2. *The Journal of General Virology*, *16*(1), 99–102.

Geisberg, J. V, Lee, W. S., Berk, A. J., & Ricciardi, R. P. (1994). The zinc finger region of the adenovirus E1A transactivating domain complexes with the TATA box binding protein. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(7), 2488–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8146144

Ghebremedhin, B. (2014). Human adenovirus: Viral pathogen with increasing importance. *European Journal of Microbiology & Immunology*, *4*(1), 26–33.

Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, *11*(4), 355–60.

Gonzalez, R. A., & Flint, S. J. (2002). Effects of mutations in the adenoviral E1B 55-kilodalton protein coding sequence on viral late mRNA metabolism. *Journal of Virology*, *76*(9), 4507–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11932416

Gooding, L. R., Elmore, L. W., Tollefson, A. E., Brady, H. A., & Wold, W. S. (1988). A 14,700 MW protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell*, *53*(3), 341–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3365766

Gooding, L. R., Sofola, I. O., Tollefson, A. E., Duerksen-Hughes, P., & Wold, W. S. (1990). The adenovirus E3-14.7K protein is a general inhibitor of tumor necrosis factor-mediated cytolysis. *Journal of Immunology*, *145*(9), 3080–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2145367

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., & Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research*, *38*(Web Server issue), W695-9.

Green, M., Wold, W. S., Brackmann, K., & Cartas, M. A. (1980). Studies on early proteins and transformation proteins of human adenoviruses. *Cold Spring Harbor Symposia on Quantitative Biology*, *44 Pt 1,* 457–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6933042

Guo, R., Zheng, L., Park, J. W., Lv, R., Chen, H., Jiao, F., … Shi, Y. (2014). BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Molecular Cell*, *56*(2), 298–310.

Harter, M. R., Liu, C.-D., Shen, C.-L., Gonzalez-Hurtado, E., Zhang, Z.-M., Xu, M., … Song, J. (2016). BS69/ZMYND11 C-Terminal Domains Bind and Inhibit EBNA2. *PLoS Pathogens*, *12*(2), e1005414.

Hateboer, G., Gennissen, A., Ramos, Y. F., Kerkhoven, R. M., Sonntag-Buck, V., Stunnenberg, H. G., & Bernards, R. (1995). BS69, a novel adenovirus E1A-associated protein that inhibits E1A transactivation. *The EMBO Journal*, *14*(13), 3159–69. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=394377&tool=pmcentrez&rendertype=abstract

Hawkins, L. K., & Wold, W. S. (1992). A 12,500 MW protein is coded by region E3 of adenovirus. *Virology*, *188*(2), 486–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1585632

Hearing, P., & Shenk, T. (1983). The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell*, *33*(3), 695–703. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6871991

Hilleman, M. R., & Werner, J. H. (1954). Recovery of new agent from patients with acute respiratory illness. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine*, *85*(1), 183–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13134329

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, *77*(1), 51–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2744487

Huang, H., Li, L., Wu, C., Schibli, D., Colwill, K., Ma, S., … Li, S. S.-C. (2008). Defining the specificity space of the human SRC homology 2 domain. *Molecular & Cellular Proteomics*, *7*(4), 768–84.

Huebner, R. J., Rowe, W. P., Ward, T. G., Parrott, R. H., & Bell, J. A. (1954). Adenoidal-pharyngeal-conjunctival agents: a newly recognized group of common viruses of the respiratory system. *The New England Journal of Medicine*, *251*(27), 1077–86.

Isobe, T., Uchida, C., Hattori, T., Kitagawa, K., Oda, T., & Kitagawa, M. (2006). Ubiquitin-dependent degradation of adenovirus E1A protein is inhibited by BS69. *Biochemical and Biophysical Research Communications*, *339*(1), 367–74.

Jones, M. S., Harrach, B., Ganac, R. D., Gozum, M. M. A., Dela Cruz, W. P., Riedel, B., … Schnurr, D. P. (2007). New adenovirus species found in a patient presenting with gastroenteritis. *Journal of Virology*, *81*(11), 5978–84.

Jones, N., & Shenk, T. (1979). An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(8), 3665–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/291030

Kateb, F., Perrin, H., Tripsianes, K., Zou, P., Spadaccini, R., Bottomley, M., … Sattler, M. (2013). Structural and functional analysis of the DEAF-1 and BS69 MYND domains. *PloS One*, *8*(1), e54715.

Kawata, H., Yamada, K., Shou, Z., Mizutani, T., Yazawa, T., Yoshino, M., … Miyamoto, K. (2003). Zinc-fingers and homeoboxes (ZHX) 2, a novel member of the ZHX family, functions as a transcriptional repressor. *Biochemical Journal*, *373*(3), 747–757.

Kim, H.-J., & Foster, M. P. (2002). Characterization of Ad5 E3-14.7K, an adenoviral

inhibitor of apoptosis: structure, oligomeric state, and metal binding. *Protein Science : A Publication of the Protein Society*, *11*(5), 1117–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11967368

King, C. R., Cohen, M. J., Fonseca, G. J., Dirk, B. S., Dikeakos, J. D., & Mymryk, J. S. (2016). Functional and Structural Mimicry of Cellular Protein Kinase A Anchoring Proteins by a Viral Oncoprotein. *PLoS Pathogens*, *12*(5), e1005621.

Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Vajda, S. (2013). How good is automated protein docking? *Proteins*, *81*(12), 2159–66.

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., … Vajda, S. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols*, *12*(2), 255–278.

Krilov, L. R. (2005). Adenovirus infections in the immunocompromised host. *The Pediatric Infectious Disease Journal*, *24*(6), 555–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15933569

Kurozumi, K., Nishita, M., Yamaguchi, K., Fujita, T., Ueno, N., & Shibuya, H. (1998). BRAM1, a BMP receptor-associated molecule involved in BMP signalling. *Genes to Cells*, *3*(4), 257–64.

Ladendorff, N. E., Wu, S., & Lipsick, J. S. (2001). BS69, an adenovirus E1A-associated protein, inhibits the transcriptional activity of c-Myb. *Oncogene*, *20*(1), 125–32.

Lang, S. E., & Hearing, P. (2003). The adenovirus E1A oncoprotein recruits the cellular TRRAP/GCN5 histone acetyltransferase complex. *Oncogene*, *22*(18), 2836–41.

Lau, L., Gray, E. E., Brunette, R. L., & Stetson, D. B. (2015). DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science*, *350*(6260), 568–71.

Lewis, P. F., Schmidt, M. A., Lu, X., Erdman, D. D., Campbell, M., Thomas, A., … Gilbert, D. (2009). A community-based outbreak of severe respiratory illness caused

by human adenovirus serotype 14. *The Journal of Infectious Diseases*, *199*(10), 1427–34.

Li, X., Romero, P., Rani, M., Dunker, A. K., & Obradovic, Z. (1999). Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome Informatics. Workshop on Genome Informatics*, *10*, 30–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11072340

Li, Y., Kang, J., Friedman, J., Tarassishin, L., Ye, J., Kovalenko, A., … Horwitz, M. S. (1999). Identification of a cell protein (FIP-3) as a modulator of NF-kappaB activity and as a target of an adenovirus inhibitor of tumor necrosis factor alpha-induced apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(3), 1042–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9927690

Li, Y., Kang, J., & Horwitz, M. S. (1997). Interaction of an adenovirus 14.7-kilodalton protein inhibitor of tumor necrosis factor alpha cytolysis with a new member of the GTPase superfamily of signal transducers. *Journal of Virology*, *71*(2), 1576–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8995684

Li, Y., & Li, H. (2012). Many keys to push: diversifying the “readership” of plant homeodomain fingers. *Acta Biochimica et Biophysica Sinica*, *44*(1), 28–39.

Lichtenstein, D. L., Toth, K., Doronin, K., Tollefson, A. E., & Wold, W. S. M. (2004). Functions and mechanisms of action of the adenovirus E3 proteins. *International Reviews of Immunology*, *23*(1–2), 75–111. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14690856

Lillie, J. W., & Green, M. R. (1989). Transcription activation by the adenovirus E1a protein. *Nature*, *338*(6210), 39–44.

Lindenbaum, J. O., Field, J., & Hurwitz, J. (1986). The adenovirus DNA binding protein and adenovirus DNA polymerase interact to catalyze elongation of primed DNA templates. *The Journal of Biological Chemistry*, *261*(22), 10218–27. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/2942538

Lipinski, K. S., Esche, H., & Brockmann, D. (1998). Amino acids 1-29 of the adenovirus serotypes 12 and 2 E1A proteins interact with rap30 (TF(II)F) and TBP in vitro. *Virus Research*, *54*(1), 99–106. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9660075

Liu, F., & Green, M. R. (1994). Promoter targeting by adenovirus E1a through interaction with different cellular DNA-binding domains. *Nature*, *368*(6471), 520–5.

Liu, H., Jin, L., Koh, S. B. S., Atanasov, I., Schein, S., Wu, L., & Zhou, Z. H. (2010). Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science*, *329*(5995), 1038–43.

Liu, L., Zhen, X. T., Denton, E., Marsden, B. D., & Schapira, M. (2012). ChromoHub: a data hub for navigators of chromatin-mediated signalling. *Bioinformatics*, *28*(16), 2205–6.

Liu, Y., Chen, W., Gaudet, J., Cheney, M. D., Roudaia, L., Cierpicki, T., … Bushweller, J. H. (2007). Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer Cell*, *11*(6), 483–97.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, *25*(4), 402–8.

Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., … Richardson, D. C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*, *50*(3), 437–50.

Lowe, S. W., & Ruley, H. E. (1993). Stabilization of the p53 tumor suppressor is induced by adenovirus 5 E1A and accompanies apoptosis. *Genes & Development*, *7*(4), 535–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8384579

Lu, H., & Levine, A. J. (1995). Human TAFII31 protein is a transcriptional coactivator of

the p53 protein. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(11), 5154–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7761466

Mackey, J. K., Rigden, P. M., & Green, M. (1976). Do highly oncogenic group A human adenoviruses cause human cancer? Analysis of human tumors for adenovirus 12 transforming DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *73*(12), 4657–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1070016

Mackmull, M.-T., Iskar, M., Parca, L., Singer, S., Bork, P., Ori, A., & Beck, M. (2015). Histone Deacetylase Inhibitors (HDACi) Cause the Selective Depletion of Bromodomain Containing Proteins (BCPs). *Molecular & Cellular Proteomics*, *14*(5), 1350–60.

Madisch, I., Harste, G., Pommer, H., & Heim, A. (2005). Phylogenetic analysis of the main neutralization and hemagglutination determinants of all human adenovirus prototypes as a basis for molecular classification and taxonomy. *Journal of Virology*, *79*(24), 15265–76.

Mal, A., Poon, R. Y., Howe, P. H., Toyoshima, H., Hunter, T., & Harter, M. L. (1996). Inactivation of p27Kip1 by the viral E1A oncoprotein in TGFbeta-treated cells. *Nature*, *380*(6571), 262–5.

Martin, K. J., Lillie, J. W., & Green, M. R. (1990). Evidence for interaction of different eukaryotic transcriptional activators with distinct cellular targets. *Nature*, *346*(6280), 147–52.

Marton, M. J., Baim, S. B., Ornelles, D. A., & Shenk, T. (1990). The adenovirus E4 17-kilodalton protein complexes with the cellular transcription factor E2F, altering its DNA-binding properties and stimulating E1A-independent accumulation of E2 mRNA. *Journal of Virology*, *64*(5), 2345–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2139141

Masselink, H., & Bernards, R. (2000). The adenovirus E1A binding protein BS69 is a corepressor of transcription through recruitment of N-CoR. *Oncogene*, *19*(12), 1538–46.

Masselink, H., Vastenhouw, N., & Bernards, R. (2001). B-myb rescues ras-induced premature senescence, which requires its transactivation domain. *Cancer Letters*, *171*(1), 87–101. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11485831

Mathews, M. B., & Shenk, T. (1991). Adenovirus virus-associated RNA and translation control. *Journal of Virology*, *65*(11), 5657–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1920611

Meng, X., Yang, Y.-F., Cao, X., Govindan, M. V, Shuen, M., Hollenberg, A. N., … Walfish, P. G. (2003). Cellular context of coregulator and adaptor proteins regulates human adenovirus 5 early region 1A-dependent gene activation by the thyroid hormone receptor. *Molecular Endocrinology*, *17*(6), 1095–105.

Moise, A. R., Grant, J. R., Vitalis, T. Z., & Jefferies, W. A. (2002). Adenovirus E3-6.7K maintains calcium homeostasis and prevents apoptosis and arachidonic acid release. *Journal of Virology*, *76*(4), 1578–87. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11799152

Moskowitz, A. M., Belnap, N., Siniard, A. L., Szelinger, S., Claasen, A. M., Richholt, R. F., … Schrauwen, I. (2016). A de novo missense mutation in ZMYND11 is associated with global developmental delay, seizures, and hypotonia. *Cold Spring Harbor Molecular Case Studies*, *2*(5), a000851.

Neduva, V., & Russell, R. B. (2005). Linear motifs: evolutionary interaction switches. *FEBS Letters*, *579*(15), 3342–5.

Neuhoff, V., Arold, N., Taube, D., & Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis*, *9*(6), 255–62.

Nevins, J. R., Ginsberg, H. S., Blanchard, J. M., Wilson, M. C., & Darnell, J. E. (1979). Regulation of the primary expression of the early adenovirus transcription units. *Journal of Virology*, *32*(3), 727–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/513202

O'Shea, C., Klupsch, K., Choi, S., Bagus, B., Soria, C., Shen, J., … Stokoe, D. (2005). Adenoviral proteins mimic nutrient/growth signals to activate the mTOR pathway for viral replication. *The EMBO Journal*, *24*(6), 1211–21.

Obert, S., O'Connor, R. J., Schmid, S., & Hearing, P. (1994). The adenovirus E4-6/7 protein transactivates the E2 promoter by inducing dimerization of a heteromeric E2F complex. *Molecular and Cellular Biology*, *14*(2), 1333–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8289811

Ogata-Kawata, H., Yamada, K., Uesaka-Yoshino, M., Kagawa, N., & Miyamoto, K. (2007). BS69, a corepressor interacting with ZHX1, is a bifunctional transcription factor. *Frontiers in Bioscience*, *12*, 1911–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17127430

Ostapchuk, P., & Hearing, P. (2008). Adenovirus IVa2 protein binds ATP. *Journal of Virology*, *82*(20), 10290–4.

Papillon-Cavanagh, S., Lu, C., Gayden, T., Mikael, L. G., Bechet, D., Karamboulas, C., … Jabado, N. (2017). Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. *Nature Genetics*, *49*(2), 180–185.

Pelka, P., Ablack, J. N. G., Fonseca, G. J., Yousef, A. F., & Mymryk, J. S. (2008). Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes. *Journal of Virology*, *82*(15), 7252–63.

Perricaudet, M., Akusjärvi, G., Virtanen, A., & Pettersson, U. (1979). Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. *Nature*, *281*(5733), 694–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/551290

Plotnik, J. P., Budka, J. A., Ferris, M. W., & Hollenhorst, P. C. (2014). ETS1 is a genome-wide effector of RAS/ERK signaling in epithelial cells. *Nucleic Acids Research*, *42*(19), 11928–40.

Plotnik, J. P., & Hollenhorst, P. C. (2017). Interaction with ZMYND11 mediates opposing roles of Ras-responsive transcription factors ETS1 and ETS2. *Nucleic Acids Research*, gkx039.

Qin, S., & Min, J. (2014). Structure and function of the nucleosome-binding PWWP domain. *Trends in Biochemical Sciences*, *39*(11), 536–47.

Querido, E., Blanchette, P., Yan, Q., Kamura, T., Morrison, M., Boivin, D., … Branton, P. E. (2001). Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes & Development*, *15*(23), 3104–17.

Reddy, V. S., Natchiar, S. K., Stewart, P. L., & Nemerow, G. R. (2010). Crystal structure of human adenovirus at 3.5 A resolution. *Science*, *329*(5995), 1071–5.

Romero, P., Obradovic, Z., & Dunker, A. K. (1997). Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. *Genome Informatics. Workshop on Genome Informatics*, *8*, 110–124. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11072311

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H., & Ward, T. G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine*, *84*(3), 570–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13134217

Russell, W. C. (2009). Adenoviruses: update on structure and function. *The Journal of General Virology*, *90*(Pt 1), 1–20.

Sambursky, R. P., Fram, N., & Cohen, E. J. (2007). The prevalence of adenoviral conjunctivitis at the Wills Eye Hospital Emergency Room. *Optometry*, *78*(5), 236–9.

San Martín, C. (2012). Latest insights on adenovirus structure and assembly. *Viruses*, *4*(5), 847–77.

Sánchez-Duffhues, G., Hiepen, C., Knaus, P., & Ten Dijke, P. (2015). Bone morphogenetic protein signaling in bone homeostasis. *Bone*, *80*, 43–59.

Schindler, U., Beckmann, H., & Cashmore, A. R. (1993). HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *The Plant Journal*, *4*(1), 137–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8106082

Schwartzentruber, J., Korshunov, A., Liu, X.-Y., Jones, D. T. W., Pfaff, E., Jacob, K., … Jabado, N. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature*, *482*(7384), 226–31.

Sherr, C. J., & McCormick, F. (2002). The RB and p53 pathways in cancer. *Cancer Cell*, *2*(2), 103–12.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*, 539.

Sims, R. J., Weihe, E. K., Zhu, L., O'Malley, S., Harriss, J. V, & Gottlieb, P. D. (2002). m-Bop, a repressor protein essential for cardiogenesis, interacts with skNAC, a heart- and muscle-specific transcription factor. *The Journal of Biological Chemistry*, *277*(29), 26524–9.

Sohrab, S. S., Kamal, M. A., Atkinson, R. L., Alawi, M. M., & Azhar, E. I. (2017). Viral Infection and Obesity: current status and future prospective. *Current Drug Metabolism*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/28093994

Somasundaram, K., Jayaraman, G., Williams, T., Moran, E., Frisch, S., & Thimmapaya, B. (1996). Repression of a matrix metalloprotease gene by E1A correlates with its ability to bind to cell type-specific transcription factor AP-2. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(7), 3088–93.

Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8610173

Song, J., Zhang, Z., Hu, W., & Chen, Y. (2005). Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *The Journal of Biological Chemistry*, *280*(48), 40122–9.

Soria, C., Estermann, F. E., Espantman, K. C., & O'Shea, C. C. (2010). Heterochromatin silencing of p53 target genes by a small viral protein. *Nature*, *466*(7310), 1076–81.

Stec, I., Wright, T. J., van Ommen, G. J., de Boer, P. A., van Haeringen, A., Moorman, A. F., … den Dunnen, J. T. (1998). WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a Drosophila dysmorphy gene maps in the Wolf-Hirschhorn syndrome critical region and is fused to IgH in t(4;14) multiple myeloma. *Human Molecular Genetics*, *7*(7), 1071–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9618163

Stephens, C., & Harlow, E. (1987). Differential splicing yields novel adenovirus 5 E1A mRNAs that encode 30 kd and 35 kd proteins. *The EMBO Journal*, *6*(7), 2027–35. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2958276

Stracker, T. H., Carson, C. T., & Weitzman, M. D. (2002). Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. *Nature*, *418*(6895), 348–52.

Tarassishin, L., Szawlowski, P., Kidd, A. H., & Russell, W. C. (2000). An epitope on the adenovirus fibre tail is common to all human subgroups. *Archives of Virology*, *145*(4), 805–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10893158

Täuber, B., & Dobner, T. (2001). Adenovirus early E4 genes in viral oncogenesis. *Oncogene*, *20*(54), 7847–54.

Thai, M., Graham, N. A., Braas, D., Nehil, M., Komisopoulou, E., Kurdistani, S. K., … Christofk, H. R. (2014). Adenovirus E4ORF1-induced MYC activation promotes host cell anabolic glucose metabolism and virus replication. *Cell Metabolism*, *19*(4), 694–701.

Thomas, D. L., Schaack, J., Vogel, H., & Javier, R. (2001). Several E4 region functions influence mammary tumorigenesis by human adenovirus type 9. *Journal of Virology*, *75*(2), 557–68.

Tollefson, A. E., Scaria, A., Hermiston, T. W., Ryerse, J. S., Wold, L. J., & Wold, W. S. (1996). The adenovirus death protein (E3-11.6K) is required at very late stages of infection for efficient cell lysis and release of adenovirus from infected cells. *Journal of Virology*, *70*(4), 2296–306. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8642656

Tollefson, A. E., Scaria, A., Saha, S. K., & Wold, W. S. (1992). The 11,600-MW protein encoded by region E3 of adenovirus is expressed early but is greatly amplified at late stages of infection. *Journal of Virology*, *66*(6), 3633–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1316473

Tollefson, A. E., Ying, B., Doronin, K., Sidor, P. D., & Wold, W. S. M. (2007). Identification of a new human adenovirus protein encoded by a novel late l-strand transcription unit. *Journal of Virology*, *81*(23), 12918–26.

Trentin, J. J., Yabe, Y., & Taylor, G. (1962). The quest for human cancer viruses. *Science*, *137*(3533), 835–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13922417

Van den Elsen, P., Houweling, A., & Van der Eb, A. (1983). Expression of region E1b of human adenoviruses in the absence of region E1a is not sufficient for complete transformation. *Virology*, *128*(2), 377–90.

van Raaij, M. J., Mitraki, A., Lavigne, G., & Cusack, S. (1999). A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature*, *401*(6756), 935–8.

Van Roey, K., Uyar, B., Weatheritt, R. J., Dinkel, H., Seiler, M., Budd, A., … Davey, N. E. (2014). Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chemical Reviews*, *114*(13), 6733–78.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, *3*(7), RESEARCH0034. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12184808

Velasco, G., Grkovic, S., & Ansieau, S. (2006). New insights into BS69 functions. *The Journal of Biological Chemistry*, *281*(24), 16546–50.

Via, A., Uyar, B., Brun, C., & Zanzoni, A. (2015). How pathogens use linear motifs to perturb host cell networks. *Trends in Biochemical Sciences*, *40*(1), 36–48.

von der Haar, T. (2007). Optimized protein extraction for quantitative proteomics of yeasts. *PloS One*, *2*(10), e1078.

Walls, T., Shankar, A. G., & Shingadia, D. (2003). Adenovirus: an increasingly important pathogen in paediatric bone marrow transplant patients. *The Lancet. Infectious Diseases*, *3*(2), 79–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12560192

Walsh, M. P., Seto, J., Liu, E. B., Dehghan, S., Hudson, N. R., Lukashev, A. N., … Seto, D. (2011). Computational analysis of two species C human adenoviruses provides evidence of a novel virus. *Journal of Clinical Microbiology*, *49*(10), 3482–90.

Walter, M. J., Payton, J. E., Ries, R. E., Shannon, W. D., Deshmukh, H., Zhao, Y., … Ley, T. J. (2009). Acquired copy number alterations in adult acute myeloid leukemia genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 12950–5.

Wang, G., & Berk, A. J. (2002). In vivo association of adenovirus large E1A protein with the human mediator complex in adenovirus-infected and -transformed cells. *Journal of Virology*, *76*(18), 9186–93. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12186902

Wang, J., Qin, S., Li, F., Li, S., Zhang, W., Peng, J., … Shi, Y. (2014). Crystal structure of human BS69 Bromo-ZnF-PWWP reveals its role in H3K36me3 nucleosome

binding. *Cell Research*, *24*(7), 890–3.

Webb, B., & Sali, A. (2016). Comparative Protein Structure Modeling Using MODELLER. *Current Protocols in Protein Science*, *86*, 2.9.1-2.9.37.

Webster, L. C., & Ricciardi, R. P. (1991). trans-dominant mutants of E1A provide genetic evidence that the zinc finger of the trans-activating domain binds a transcription factor. *Molecular and Cellular Biology*, *11*(9), 4287–96. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1831535

Wei, G., Schaffner, A. E., Baker, K. M., Mansky, K. C., & Ostrowski, M. C. (2003). Ets-2 interacts with co-repressor BS69 to repress target gene expression. *Anticancer Research*, *23*(3A), 2173–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12894593

Wen, H., Li, Y., Xi, Y., Jiang, S., Stratton, S., Peng, D., … Shi, X. (2014). ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression. *Nature*, *508*(7495), 263–8.

White, E. (2001). Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene*, *20*(54), 7836–46.

Wickham, T. J., Mathias, P., Cheresh, D. A., & Nemerow, G. R. (1993). Integrins alpha v beta 3 and alpha v beta 5 promote adenovirus internalization but not virus attachment. *Cell*, *73*(2), 309–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8477447

Wold, W. S., Cladaras, C., Deutscher, S. L., & Kapoor, Q. S. (1985). The 19-kDa glycoprotein coded by region E3 of adenovirus. Purification, characterization, and structural analysis. *The Journal of Biological Chemistry*, *260*(4), 2424–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3882694

Yang, H., Zhang, C., Zhao, X., Wu, Q., Fu, X., Yu, B., … Huang, X. (2010). Analysis of copy number variations of BS69 in multiple types of hematological malignancies. *Annals of Hematology*, *89*(10), 959–64.

Yee, A. A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., … Arrowsmith, C. H. (2005). NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *Journal of the American Chemical Society*, *127*(47), 16512–7.

Yousef, A. F., Fonseca, G. J., Pelka, P., Ablack, J. N. G., Walsh, C., Dick, F. A., … Mymryk, J. S. (2010). Identification of a molecular recognition feature in the E1A oncoprotein that binds the SUMO conjugase UBC9 and likely interferes with polySUMOylation. *Oncogene*, *29*(33), 4693–704.

Zamanian, M., & La Thangue, N. B. (1992). Adenovirus E1a prevents the retinoblastoma gene product from repressing the activity of a cellular transcription factor. *The EMBO Journal*, *11*(7), 2603–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1385776

Zhang, Z., Smith, M. M., & Mymryk, J. S. (2001). Interaction of the E1A oncoprotein with Yak1p, a novel regulator of yeast pseudohyphal differentiation, and related mammalian kinases. *Molecular Biology of the Cell*, *12*(3), 699–710. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11251081

Curriculum Vitae

Name: Ali Zhang

Post-secondary Education and Degrees:
The University of Western Ontario
London, Ontario, Canada
2011-2015 H.B.MSc.

Honours and Awards:
NSERC Canada Graduate Scholarships – Master's
2016-2017

CIHR Strategic Training Program in Cancer Research and Technology Transfer (CaRTT) Trainee Award
2015-2016

University of Western Ontario Continuing Admission Scholarship
2011-2015

CaRTT Summer Studentship Award
2014

Microbiology & Immunology Summer Undergraduate Research Fellowship
2014

Related Work Experience
Teaching Assistant
The University of Western Ontario
2016

Publication:

King, C. R., **Zhang, A.**, & Mymryk, J. S. (2016). The Persistent Mystery of Adenovirus Persistence. *Trends in Microbiology*, *24*(5), 323–4.

Presentations:

Poster	2016 Infection and Immunity Research Forum, London ON
Podium	2016 DNA Tumour Virus Meeting, Montreal QC
Poster	2016 Oncology Research and Education Day, London ON
Poster	2016 London Health Research Day, London ON
Podium	2015 London Cancer Trainee Forum, London ON
Poster	2015 Infection and Immunity Research Forum, London ON
Poster	2015 Oncology Research and Education Day, London ON